



```
CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM
CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)
CC are disclosed, as are marker oligonucleotide sequences: tumour
CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal
CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers
CC (PEM) ABL91903-ABL91995
XX ABL91903-ABL91995
SQ Sequence 500 AA;

Query Match 100.0%; Score 2691; DB 5; Length 500;
Best Local Similarity 100.0%; Pred. No. 3.2e-249;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRGELWLLVLVLRRAALSPQAGHDEGPGSGWAAKGTVRGNRRARSPGHVSPDR 60
Db 1 MRGELWLLVLVLRRAALSPQAGHDEGPGSGWAAKGTVRGNRRARSPGHVSPDR 60

Qy 61 TQLSODLGGGTLMADTLPDNRTRVEDNHSYVSRLYGPPSEPHSRELWVDVAEANRSQVK 120
Db 61 TQLSODLGGGTLMADTLPDNRTRVEDNHSYVSRLYGPPSEPHSRELWVDVAEANRSQVK 120

Qy 121 IHTILSNTHRQASRVLSFDFPFYGHPLRQITATGGFIFMGDVIHRMLTATQVAPLMA 180
Db 121 IHTILSNTHRQASRVLSFDFPFYGHPLRQITATGGFIFMGDVIHRMLTATQVAPLMA 180

Qy 181 NFNPGYSDNSTVYVFDNGTVFVQWHDVYLQGWEDKGSFTFQAAALHHDGRIVFAYKEIPM 240
Db 181 NFNPGYSDNSTVYVFDNGTVFVQWHDVYLQGWEDKGSFTFQAAALHHDGRIVFAYKEIPM 240

Qy 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPESSRRSIFEVHRIELDPKVTSMASVEFTP 300
Db 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPESSRRSIFEVHRIELDPKVTSMASVEFTP 300

Qy 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLRQCSGFDYRQEWMDYGCQAEGRMCDQFQ 360
Db 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLRQCSGFDYRQEWMDYGCQAEGRMCDQFQ 360

Qy 361 DEHDASPTSPSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPKTKGT 420
Db 361 DEHDASPTSPSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPKTKGT 420

Qy 421 PVHLGTTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHHPAMKFRSHDPHST 480
Db 421 PVHLGTTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHHPAMKFRSHDPHST 480

RESULT 2
ABUS4456
ID ABUS4456 standard; protein; 500 AA.
XX
AC ABUS4456;
XX
DT 12-MAR-2003 (first entry)
XX
DE Human tumour endothelial marker TEM 17.
XX
KW Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;
KW Tumour endothelial marker; normal endothelial marker; PEM;
KW pan-endothelial marker; polycystic kidney disease; psoriasis;
KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;
KW neovascularization; immune response; cytostatic; antidiabetic;
KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic.
XX
OS Homo sapiens.
XX
PN WO200283874-A2.
XX
PD 24-OCT-2002.
XX
```

```
PF 10-APR-2002; 2002WO-US008253.
XX
PR 11-APR-2001; 2001US-0282850P.
PR 06-FEB-2002; 2002US-0354262P.
XX
PA (UYJO ) UNIV JOHNS HOPKINS.
XX
PI Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;
XX WPI; 2003-093016/08.
DR N-PSDB; ABX72028.
XX
PT New purified human transmembrane protein, designated as tumor endothelial
PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,
PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or
PT psoriasis.
XX
PS Disclosure; Page 221-222; 374pp; English.
XX
CC The present invention relates to a novel method for the isolation of
CC endothelial cells (ECs), and the identification of genes expressed in
CC normal and tumour ECs. Tumour endothelial marker (TEM), normal
CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are
CC identified in human ECs. The human EC marker proteins and the
CC polynucleotide sequences encoding them are useful for detecting,
CC diagnosing or treating tumours as well as polycystic kidney disease,
CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also
CC useful for inhibiting neovascularization or tumour angiogenesis, for
CC inducing an immune response to tumour endothelial cells in a patient, or
CC for identifying candidate drugs for treating tumours. The present
CC sequence represents a human TEM or NEM protein of the invention
XX
SQ Sequence 500 AA;

Query Match 100.0%; Score 2691; DB 6; Length 500;
Best Local Similarity 100.0%; Pred. No. 3.2e-249;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRGELWLLVLVLRRAALSPQAGHDEGPGSGWAAKGTVRGNRRARSPGHVSPDR 60
Db 1 MRGELWLLVLVLRRAALSPQAGHDEGPGSGWAAKGTVRGNRRARSPGHVSPDR 60

Qy 61 TQLSODLGGGTLMADTLPDNRTRVEDNHSYVSRLYGPPSEPHSRELWVDVAEANRSQVK 120
Db 61 TQLSODLGGGTLMADTLPDNRTRVEDNHSYVSRLYGPPSEPHSRELWVDVAEANRSQVK 120

Qy 121 IHTILSNTHRQASRVLSFDFPFYGHPLRQITATGGFIFMGDVIHRMLTATQVAPLMA 180
Db 121 IHTILSNTHRQASRVLSFDFPFYGHPLRQITATGGFIFMGDVIHRMLTATQVAPLMA 180

Qy 181 NFNPGYSDNSTVYVFDNGTVFVQWHDVYLQGWEDKGSFTFQAAALHHDGRIVFAYKEIPM 240
Db 181 NFNPGYSDNSTVYVFDNGTVFVQWHDVYLQGWEDKGSFTFQAAALHHDGRIVFAYKEIPM 240

Qy 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPESSRRSIFEVHRIELDPKVTSMASVEFTP 300
Db 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPESSRRSIFEVHRIELDPKVTSMASVEFTP 300

Qy 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLRQCSGFDYRQEWMDYGCQAEGRMCDQFQ 360
Db 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLRQCSGFDYRQEWMDYGCQAEGRMCDQFQ 360

Qy 361 DEHDASPTSPSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPKTKGT 420
Db 361 DEHDASPTSPSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPKTKGT 420

Qy 421 PVHLGTTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHHPAMKFRSHDPHST 480
Db 421 PVHLGTTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHHPAMKFRSHDPHST 480

YAEVPSGHEKEGFMEEAQC 500
YAEVPSGHEKEGFMEEAQC 500
```



CC cytostatic compounds through the regulation of the expression of a gene  
CC or activity of a protein associated with malignant neoplasia. The method  
CC is useful for prediction, diagnosis or prognosis of malignant neoplasia  
CC such as breast cancer, ovarian cancer, gastric cancer, colon cancer,  
CC esophageal cancer, mesenchymal cancer, bladder cancer or non-small cell  
CC lung cancer. The polynucleotides and polypeptides defined in the  
CC specification, antisense polynucleotides targeting the polynucleotides,  
CC antibodies targeting either one of the polynucleotides or polypeptides,  
CC and compounds identified by the screening methods are useful for  
CC preventing or treating malignant neoplasia. The disease treated is  
CC preferably breast cancer. The present sequence is that of a human  
CC malignant neoplasia-related protein which may be used in the method of  
CC the invention.

XX  
SQ Sequence 500 AA;  
Query Match 100.0%; Score 2691; DB 8; Length 500;  
Best Local Similarity 100.0%; Pred. No. 3.2e-249;  
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRGELMLLVLRRAARALSPQAGHDEGPGSGAAKGTVRGNRRRARESPGHVSEPD 60  
Db 1 MRGELMLLVLRRAARALSPQAGHDEGPGSGAAKGTVRGNRRRARESPGHVSEPD 60  
Qy 61 TQSLDGLGGTTLAMDITLPDNRTRVEDNHSYYSRLYGPSEPHSRELWVDVAENRSQVK 120  
Db 61 TQSLDGLGGTTLAMDITLPDNRTRVEDNHSYYSRLYGPSEPHSRELWVDVAENRSQVK 120  
Qy 121 IHTILSNTHRQASRVLSFDFFPGHPLRQITATGTFIPMGDVIHRLMTATQVAPLMA 180  
Db 121 IHTILSNTHRQASRVLSFDFFPGHPLRQITATGTFIPMGDVIHRLMTATQVAPLMA 180  
Qy 181 NFNPYSDNSTVYFDNGTVFVQWDHVLQGWEDKGSFTFQAALHHDGRIVFAYKEIPM 240  
Db 181 NFNPYSDNSTVYFDNGTVFVQWDHVLQGWEDKGSFTFQAALHHDGRIVFAYKEIPM 240  
Qy 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPSRRRSIFEYHRIELDSPKTSMSAVEFTP 300  
Db 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPSRRRSIFEYHRIELDSPKTSMSAVEFTP 300  
Qy 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLQRCSSGFDYRQEWMDYGCQAEGRMCDPQ 360  
Db 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLQRCSSGFDYRQEWMDYGCQAEGRMCDPQ 360  
Qy 361 DEHDSASPTSFSPYDGLTITSSSLFIDSLTTEDDTKLNYPAGDGLQNNLSPTKGT 420  
Db 361 DEHDSASPTSFSPYDGLTITSSSLFIDSLTTEDDTKLNYPAGDGLQNNLSPTKGT 420  
Qy 421 FVHLGTTIGVILVALLVAAILAGIYINGHPTSNAAALFFIERRPHHPAMKFRSHPDHST 480  
Db 421 FVHLGTTIGVILVALLVAAILAGIYINGHPTSNAAALFFIERRPHHPAMKFRSHPDHST 480  
Qy 481 YAEVEPSGHEKEGFMPEAEQC 500  
Db 481 YAEVEPSGHEKEGFMPEAEQC 500

RESULT 5  
ADI21554  
ID ADI21554 standard; protein; 527 AA.  
XX  
AC ADI21554;  
XX  
DT 15-APR-2004 (first entry)  
XX  
DE Novel human polypeptide #33.  
XX forensic; nutritional source; damaged tissue; diseased tissue;  
KW myeloid cell disorder; lymphoid cell disorder;  
KW bone cartilage tissue growth; tendon tissue growth;  
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;  
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.  
XX

OS Homo sapiens.  
XX  
PN WO2003025148-A2.  
XX  
PD 27-MAR-2003.  
XX  
PF 19-SEP-2002; 2002WO-US029964.  
XX  
PR 19-SEP-2001; 2001US-0323739P.  
PR 13-SEP-2002; 2002US-00323739.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;  
PI Haley-Vicente D;  
XX  
DR WPI: 2003-354603/33.  
DR N-PSDB; ADI21334.  
XX  
XX New polynucleotides and secreted proteins, useful for treating myeloid or  
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve  
PT tissue growth or regeneration, in wound healing, and in tissue repair and  
PT replacement.  
XX  
PS Example 3; SEQ ID NO 805; 156pp; English.  
XX  
CC The invention relates to an isolated polynucleotide encoding a  
CC polypeptide with biological activity. The polynucleotides and  
CC polypeptides are useful in diagnostics, forensics, gene mapping,  
CC identification of mutations responsible for genetic disorders and other  
CC traits, to assess biodiversity, as nutritional sources or supplements.  
CC The polynucleotides may also be used as molecular weight markers,  
CC chromosome markers or map related gene positions, or as an antigen to  
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are  
CC useful for raising antibodies, as markers for tissues in which the  
CC corresponding polypeptide is expressed, for re-engineering damaged or  
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in  
CC bone cartilage, tendon, ligament and/or nerve tissue growth or  
CC regeneration, in wound healing, in tissue repair and replacement, in  
CC healing of burns, incisions and ulcers, and in treating cancer. The  
CC present sequence represents the amino acid sequence of a novel human  
CC polypeptide.  
XX  
SQ Sequence 527 AA;  
Query Match 100.0%; Score 2691; DB 7; Length 527;  
Best Local Similarity 100.0%; Pred. No. 3.5e-249;  
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRGELMLLVLRRAARALSPQAGHDEGPGSGAAKGTVRGNRRRARESPGHVSEPD 60  
Db 28 MRGELMLLVLRRAARALSPQAGHDEGPGSGAAKGTVRGNRRRARESPGHVSEPD 87  
Qy 61 TQSLDGLGGTTLAMDITLPDNRTRVEDNHSYYSRLYGPSEPHSRELWVDVAENRSQVK 120  
Db 88 TQSLDGLGGTTLAMDITLPDNRTRVEDNHSYYSRLYGPSEPHSRELWVDVAENRSQVK 147  
Qy 121 IHTILSNTHRQASRVLSFDFFPGHPLRQITATGTFIPMGDVIHRLMTATQVAPLMA 180  
Db 148 IHTILSNTHRQASRVLSFDFFPGHPLRQITATGTFIPMGDVIHRLMTATQVAPLMA 207  
Qy 181 NFNPYSDNSTVYFDNGTVFVQWDHVLQGWEDKGSFTFQAALHHDGRIVFAYKEIPM 240  
Db 208 NFNPYSDNSTVYFDNGTVFVQWDHVLQGWEDKGSFTFQAALHHDGRIVFAYKEIPM 267  
Qy 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPSRRRSIFEYHRIELDSPKTSMSAVEFTP 300  
Db 268 SVPEISSQHPVKTGLSDAFMILNPSDPVPSRRRSIFEYHRIELDSPKTSMSAVEFTP 327  
Qy 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLQRCSSGFDYRQEWMDYGCQAEGRMCDPQ 360  
Db 328 LPTCLQHRSCDACMSSDLTFNCSCWCHVLQRCSSGFDYRQEWMDYGCQAEGRMCDPQ 387

QY 361 DEDHDSASPDTSFSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGGLQNNLSPKTKGT 420  
 Db 388 DEDHDSASPDTSFSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGGLQNNLSPKTKGT 447  
 QY 421 PVHLGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHST 480  
 Db 448 PVHLGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHST 507  
 QY 481 YAEVPSGHEKEGFMEAEQC 500  
 Db 508 YAEVPSGHEKEGFMEAEQC 527

RESULT 6  
 ADI21553  
 ID ADI21553 standard; protein; 527 AA.  
 AC ADI21553;  
 XX  
 XX 15-APR-2004 (first entry)  
 XX  
 XX Novel human polypeptide #32.  
 DT  
 DE  
 KW forensic; nutritional source; damaged tissue; diseased tissue;  
 KW myeloid cell disorder; lymphoid cell disorder;  
 KW bone cartilage tissue growth; tendon tissue growth;  
 KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;  
 KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.  
 OS Homo sapiens.  
 XX  
 XX W02003025148-A2.  
 XX  
 XX 27-MAR-2003.  
 XX  
 XX 19-SEP-2002; 2002WO-US029964.  
 XX  
 XX 19-SEP-2001; 2001US-0323739P.  
 PR 13-SEP-2002; 2002US-00323739.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX  
 XX Tang YT, Auendi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;  
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;  
 PI Haley-Vicente D;  
 XX  
 XX WPI; 2003-354603/33.  
 DR N-PSDB; ADI21333.  
 DR  
 XX  
 XX New polynucleotides and secreted proteins, useful for treating myeloid or  
 PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve  
 PT tissue growth or regeneration, in wound healing, and in tissue repair and  
 PT replacement.  
 XX  
 XX Example 3; SEQ ID NO 804; 156pp; English.  
 PS  
 XX The invention relates to an isolated polynucleotide encoding a  
 CC polypeptide with biological activity. The polynucleotides and  
 CC polypeptides are useful in diagnostics, forensics, gene mapping,  
 CC identification of mutations responsible for genetic disorders and other  
 CC traits, to assess biodiversity, as nutritional sources or supplements.  
 CC The polynucleotides may also be used as molecular weight markers,  
 CC chromosome markers or map related gene positions, or as an antigen to  
 CC raise anti-DNA antibodies or elicit immune response. The polypeptides are  
 CC useful for raising antibodies, as markers for tissues in which the  
 CC corresponding polypeptide is expressed, for re-engineering damaged or  
 CC diseased tissues, for treating myeloid or lymphoid cell disorders, in  
 CC bone cartilage, tendon, ligament and/or nerve tissue growth or  
 CC regeneration, in wound healing, in tissue repair and replacement, in  
 CC healing of burns, incisions and ulcers, and in treating cancer. The  
 CC present sequence represents the amino acid sequence of a novel human  
 CC polypeptide.

XX SQ Sequence 527 AA;  
 Query Match 100.0%; Score 2691; DB 7; Length 527;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-249;  
 Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRGELMLLVLAAREALSPQAGHDEGPGSGAAKGTVRGNRRARESPGHVSEDR 60  
 Db 28 MRGELMLLVLAAREALSPQAGHDEGPGSGAAKGTVRGNRRARESPGHVSEDR 87  
 QY 61 TQLSQDLGGGTGLAMDITLQNRTRVVEDNHSYVYVSRSLYGPSEPHSRELWVDVAENRSQVK 120  
 Db 88 TQLSQDLGGGTGLAMDITLQNRTRVVEDNHSYVYVSRSLYGPSEPHSRELWVDVAENRSQVK 147  
 QY 121 IHTILSNTHRQASRVLSFDPFPGHPLRQITTIATGFIIMGDVIHRLMTATQVAPLMA 180  
 Db 148 IHTILSNTHRQASRVLSFDPFPGHPLRQITTIATGFIIMGDVIHRLMTATQVAPLMA 207  
 QY 181 NFNPGYSDNSTVYFQNGTVFVQWDHVLQGWEDKGSFTFOALHHDGRIVFAYKEIPM 240  
 Db 208 NFNPGYSDNSTVYFQNGTVFVQWDHVLQGWEDKGSFTFOALHHDGRIVFAYKEIPM 267  
 QY 241 SVPEISSSQHPVKTGLSDAFMILNPSDPVPESSRRSIFFEYHRIELDPKVTSMGAVFTPT 300  
 Db 268 SVPEISSSQHPVKTGLSDAFMILNPSDPVPESSRRSIFFEYHRIELDPKVTSMGAVFTPT 327  
 QY 301 LPTCLQHRSCDACMSDLTFCNSCHVLCQRCSSGFDRYQEWMDYGCQAABGRCMCDFFQ 360  
 Db 328 LPTCLQHRSCDACMSDLTFCNSCHVLCQRCSSGFDRYQEWMDYGCQAABGRCMCDFFQ 387  
 QY 361 DEDHDSASPDTSFSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGGLQNNLSPKTKGT 420  
 Db 388 DEDHDSASPDTSFSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGGLQNNLSPKTKGT 447  
 QY 421 PVHLGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHST 480  
 Db 448 PVHLGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHST 507  
 QY 481 YAEVPSGHEKEGFMEAEQC 500  
 Db 508 YAEVPSGHEKEGFMEAEQC 527

RESULT 7  
 ABB90723  
 ID ABB90723 standard; protein; 1002 AA.  
 XX  
 XX ABB90723;  
 XX  
 XX 30-MAY-2002 (first entry)  
 DT Human Tumour Endothelial Marker polypeptide SEQ ID NO 179.  
 XX  
 XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;  
 KW normal endothelial marker; pan-endothelial marker; immunostimulant;  
 KW antiangiogenic; tumour; neoangiogenesis; vascularised tumour;  
 KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;  
 KW psoriasis.  
 XX  
 XX Homo sapiens.  
 XX  
 XX W0200210217-A2.  
 PN  
 XX  
 XX 07-FEB-2002.  
 PD  
 XX  
 XX 01-AUG-2001; 2001WO-US024031.  
 XX  
 XX 02-AUG-2000; 2000US-0222599P.  
 PR 11-AUG-2000; 2000US-0224360P.  
 PR 11-APR-2001; 2001US-0282850P.  
 XX  
 XX (UWJO ) UNIV JOHNS HOPKINS.  
 PA



Db 743 SVPEISSQHPVKTGLSDAFMILNPSDPVPSRRSIFVYHRIELDFSKVTSNAVEFTP 802  
Qy 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLCRCSSGFDYRQEWMDYGCQAEGRMCDPQ 360  
Db 803 LPTCLQHRSCDACMSSDLTFNCSCWCHVLCRCSSGFDYRQEWMDYGCQAEGRMCDPQ 862  
Qy 361 DEHDSASPTSPFYDGDLTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPKTKGT 420  
Db 863 DEHDSASPTSPFYDGDLTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPKTKGT 922  
Qy 421 PVHLGTVIGIVLAVLLVAAILAGIYINGHPTSNAAFFIERRPHHPAMKFRSHPDHST 480  
Db 923 PVHLGTVIGIVLAVLLVAAILAGIYINGHPTSNAAFFIERRPHHPAMKFRSHPDHST 982  
Qy 481 YAEVPSGHEKEGFMEEAQC 500  
Db 983 YAEVPSGHEKEGFMEEAQC 1002

## RESULT 9

ADI21064  
ID ADI21064 standard; protein; 488 AA.

XX AC ADI21064;

XX 15-APR-2004 (first entry)

XX Novel human protein #39.

XX forensic; nutritional source; damaged tissue; diseased tissue;

KW myeloid cell disorder; lymphoid cell disorder;

KW bone cartilage tissue growth; tendon tissue growth;

KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;

KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.

XX Homo sapiens.

XX W02003025148-A2.

XX 27-MAR-2003.

XX 19-SEP-2002; 2002WO-US029964.

XX 19-SEP-2001; 2001US-0323739P.

XX 13-SEP-2002; 2002US-00323739.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;

PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;

PI Haley-Vicente D;

XX WPI; 2003-354603/33.

XX N-PSDB; ADI21780.

XX New polynucleotides and secreted proteins, useful for treating myeloid or  
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve  
PT tissue growth or regeneration, in wound healing, and in tissue repair and  
PT replacement.

XX Claim 20; SEQ ID NO 315; 156pp; English.

XX The invention relates to an isolated polynucleotide encoding a  
CC polypeptide with biological activity. The polynucleotides and  
CC polypeptides are useful in diagnostics, forensics, gene mapping,  
CC identification of mutations responsible for genetic disorders and other  
CC traits, to assess biodiversity, as nutritional sources or supplements.  
CC The polynucleotides may also be used as molecular weight markers,  
CC chromosome markers or map related gene positions, or as an antigen to  
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are  
CC useful for raising antibodies, as markers for tissues in which the  
CC corresponding polypeptide is expressed, for re-engineering damaged or

CC diseased tissues, for treating myeloid or lymphoid cell disorders, in  
CC bone cartilage, tendon, ligament and/or nerve tissue growth or  
CC regeneration, in wound healing, in tissue repair and replacement, in  
CC healing of burns, incisions and ulcers, and in treating cancer. The  
CC present sequence represents the amino acid sequence of a novel human  
CC protein.

XX SQ Sequence 488 AA;

Query Match 96.9%; Score 2607; DB 7; Length 488;

Best Local Similarity 97.6%; Pred. No. 3.7e-241;

Matches 488; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

Qy 1 MRGELWLLVLRRAARALSPQAGHDEGPGSGAAKGTVRGNRRARSPGHVSPDR 60

Db 1 MRGELWLLVLRRAARALSPQAGHDEGPGSGAAKGTVRGNRRARSPGHVSPDR 60

Qy 61 TQLSQDLGGGTGLAMDTLPDNRTRVVDNHSYVYVSRLYGPSEPHSRELVDVAEANSQVK 120

Db 61 TQLSQDLGGGTGLAMDTLPDNRTRVVDNHSYVYVSRLYGPSEPHSRELVDVAEANSQVK 120

Qy 121 IHTILSNTHRQASRVVLSDFPFYGHPLRQITATGGFIFMGDVIIHRMLTATQYVAELMA 180

Db 121 IHTILSNTHRQASRVVLSDFPFYGHPLRQITATGGFIFMGDVIIHRMLTATQYVAELMA 180

Qy 181 NFNPYSDNSTVYVFDNGTVVQWDHVLQGWEDKGSFTFOALHHDGRIVFAYKEIPM 240

Db 181 NFNPYSDNSTVYVFDNGTVVQWDHVLQGWEDKGSFTFOALHHDGRIVFAYKEIPM 240

Qy 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPSRRSIFVYHRIELDFSKVTSNAVEFTP 300

Db 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPSRRSIFVYHRIELDFSKVTSNAVEFTP 300

Qy 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLCRCSSGFDYRQEWMDYGCQAEGRMCDPQ 360

Db 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLCRCSSGFDYRQEWMDYGCQAEGRMCDPQ 360

Qy 361 DEHDSASPTSPFYDGDLTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPKTKGT 420

Db 361 DEHDSASPTSPFYDGDLTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPKTKGT 408

Qy 421 PVHLGTVIGIVLAVLLVAAILAGIYINGHPTSNAAFFIERRPHHPAMKFRSHPDHST 480

Db 409 PVHLGTVIGIVLAVLLVAAILAGIYINGHPTSNAAFFIERRPHHPAMKFRSHPDHST 468

Qy 481 YAEVPSGHEKEGFMEEAQC 500

Db 469 YAEVPSGHEKEGFMEEAQC 488

## RESULT 10

ABB90783

ID ABB90783 standard; protein; 500 AA.

XX AC ABB90783;

XX 30-MAY-2002 (first entry)

XX Mouse Tumour Endothelial Marker polypeptide SEQ ID NO 297.

XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;  
KW normal endothelial marker; pan-endothelial marker; immunostimulant;  
KW antiangiogenic; tumour; neovascularisation; vascularised tumour;  
KW polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;  
KW psoriasis.

XX Mus musculus.

XX W0200210217-A2.

XX 07-FEB-2002.

XX 01-AUG-2001; 2001WO-US024031.

XX  
PR PR 02-AUG-2000; 2000US-0222599P.  
PR PR 11-AUG-2000; 2000US-0224360P.  
XX XX 11-APR-2001; 2001US-0282850P.  
PA (UYJO ) UNIV JOHNS HOPKINS.  
XX St Croix B, Kinzler KW, Vogelstein B;  
XX WPI; 2002-291856/33.  
DR N-PSDB; ABL921136.  
XX An isolated molecule comprising an antibody variable region which specifically binds to an extracellular domain of a tumor endothelial marker (TEM) protein, useful for inhibiting tumor growth.  
XX Disclosure; Page 301-302; 331pp; English.  
XX The invention relates to an isolated molecule comprising an antibody variable region which specifically binds to an extracellular domain of a tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740, ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM proteins have cytostatic, immunostimulant and neoangiogenic activity. They are useful for inhibiting tumour growth, neoangiogenesis in subjects bearing a vascularised tumour, polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789) are disclosed, as are marker oligonucleotide sequences: tumour endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers (PEM) ABL91903-ABL91995

XX Sequence 500 AA;

Query Match 82.1%; Score 2209; DB 5; Length 500;  
Best Local Similarity 81.6%; Pred. No. 7e-203;  
Matches 409; Conservative 46; Mismatches 44; Indels 2; Gaps 2;

QY 1 MRGELMLL-VLVLRRAARALSPQGAGHDGPSCGAAKTGTVRGNRRARPSPGHVSEPD 59  
Db 1 MRAQLWLLQLLLRGARALSPATPAGNEGDSAWTAKTRTQGSRRPRESPAQLKPG 60

QY 60 RTOLSODLGGLTAMDTLDPNRTVVEDNHSSYYVSRSLVGPSPHRELVWDVAENRSOV 119  
Db KTLQSOLDGGSLAIDTLDPNRTVVEDNHSSYYVSRVGPGEKQSDQLWDLAVANRSHV 120

QY 120 KIHTILSNTHROASRVVLSDFPFYGHPLRQTITATGGFIHGMDLMHLMTATQVAPLM 179  
Db KIHRILSSSHROASRVVLSDFPFYGHPLRQTITATGGFIHGMDLMHLMTATQVAPLM 180

QY 180 ANFNPGYSDNSTVVFVDNGTVFVQMDHYVLQGWEDKGSTFTQAALHHDGRIVFAIKBIP 239  
Db 181 ANFNPGYSDNSTVAYFDNGTVFVQMDHYVLQDREDGTSFTQAALHHDGRIVFGYKEIP 240

QY 240 MSVPETSSSQHPVKVTGLSADFMLNPSDPVPSRRRSIFEXHRIELDPSKVTSMSAVEFT 299  
Db 241 MAVLDISSAQHPYKAGLSADFMLNSSPVSQRRTIFEXHRVELDSSKITTTSAVEFT 300

QY 300 PLPTCLQHRCSDACMSDLTFNCWCWCHVLQRCSSGFDRYRQSWMDYGCAQAEGRMCEDF 359  
Db 301 PLPTCLHQSCDTCVSSNLTFNCWCWCHVLQRCSSGFDRYRQEWLYTGCAQAEKTCEDF 360

QY 360 QDEHDDASGPDTSFSFYDGDLTTTSSSLFIDSLTDEDTKLNPYAGGDLQNLSPKTKG 419  
Db 361 QDDSHYSASPDSFSFPNGD-STTSSSLFIDSLTDEDTKLNPYAEGDGLPDHSSPKSKG 419

QY 420 TPVHLGTIGVILVALVLAAILLAGIYINGHPTSNAALEFFERRPHWPMPKERSHPDHS 479  
Db 420 PPVHLGTIGVILVALVLAAILLAGIYISGHFNNSNAALFFIERPPHHWPMPKFHNPNHS 479

QY 480 TYAEVPSGHEKGFMEABQC 500  
Db TYTEVEFPSGHEKGFVEAEOC 500







Db 121 VVFDNGTVFVQWHDVYVLOGWEDKGSFTFQALHHDGRIVFAYKEIPMSVPEISSQHP 180  
 Qy 252 VKTGLSDAFMILNPSDPVPSRRRSIFEXHRIELDPKVTSMASAVEFTPLTCLQHRSCD 311  
 Db 181 VKTGLSDAFMILNPSDPVPSRRRSIFEXHRIELDPKVTSMASAVEFTPLTCLQHRSCD 240  
 Qy 312 ACMSSDLTFNCSCWCHVLQRCSSGDFDRYRQEW-MDYGCAQAEAGRMCDQD-----ED 363  
 Db 241 ACMSSDLTFNCSCWCHVLQRCSSGDFDRYRQEW-MDYGCAQAEAGRMCDQD-----QDVRLPGMRT 294  
 Qy 364 HDSASPDTSFSPYDGLTFTSSSLFIDSLTDDTKLNPYAGDGLQNNLSPTKGTGPVH 423  
 Db 295 TTSASPDTSFSPYDGLTFTSSSLFIDSLTDDTKLNPYAGDGLQNNLSPTKGTGPVH 354  
 Qy 424 LGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHSTYAE 483  
 Db 355 LGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHSTYAE 414  
 Qy 484 VEPSGHEKEGFMEAEQC 500  
 Db 415 VEPSGHEKEGFMEAEQC 431

## RESULT 15

AB001434  
 ID AB001434 standard; protein; 431 AA.

AC AB001434;

DT 07-AUG-2003 (first entry)

XX Human tumour endothelial marker 7 precursor protein.

XX Human; gene therapy; stem cell inducer; osteoporosis; regeneration;  
 XX stem cell growth factor-like activity; leukaemia; haemophilia; allergy;  
 XX Parkinson's disease; Huntington's disease; amyotrophic lateral sclerosis;  
 XX autoimmune disease; multiple sclerosis; systemic lupus erythematosus;  
 XX graft-versus-host disease; degenerative disease; Alzheimer's disease;  
 XX osteoarthritis; anaemia; tendonitis; carpal tunnel syndrome; cancer;  
 XX tumour endothelial marker 7 precursor.

XX Homo sapiens.

XX US2003022825-A1.

XX 30-JAN-2003.

XX 24-JUL-2001; 2001US-00912935.

XX 23-DEC-2000; 2000WO-US035260.

XX (NISH/) NISHIKAWA M.

XX (LABA/) LABAT I.

XX (DRMA/) DRMANAC R T.

XX (TANG/) TANG Y T.

XX (CHAO/) CHAO C.

XX Nishikawa M, Labat I, Drmanac RT, Tang YT, Chao C;

XX WPI; 2003-456302/43.

XX New stem cell growth factor-like polypeptides and polynucleotides, useful  
 XX for treating e.g. leukemia, hemophilia, osteoporosis, osteoarthritis,  
 XX graft-versus-host disease, cancers, Alzheimer's disease, Huntington's  
 XX disease.

XX Example 4; Fig 2; 98pp; English.

XX PS

XX The invention relates to a new isolated polypeptide, which has stem cell  
 XX growth factor-like activity. The new polypeptide, the polynucleotide  
 XX encoding this polypeptide, or the agonist of the polypeptide are useful  
 XX for treating a subject in need of enhanced activity or expression of stem  
 XX cell growth factor-like polypeptide. The antagonist of the polypeptide or

CC the polynucleotide is useful for treating a subject in need to inhibit  
 CC the activity or expression of stem cell growth factor-like polypeptide.  
 CC The new polypeptide or polynucleotide is particularly useful for inducing  
 CC differentiation of embryonic and adult stem cells to give rise to  
 CC different cell types. In particular, the polypeptide or polynucleotide is  
 CC useful for treating leukaemia, haemophilia, osteoporosis, osteoarthritis,  
 CC anaemia, tendonitis, carpal tunnel syndrome, autoimmune disease (e.g.  
 CC multiple sclerosis, systemic lupus erythematosus, graft-versus-host  
 CC disease or allergies), cancers or degenerative diseases (e.g. Alzheimer's  
 CC disease, Parkinson's disease, Huntington's disease or amyotrophic lateral  
 CC sclerosis) or for generating new tissues and organs that may aid patients  
 CC in need of transplanted tissues. The polynucleotide may also be used in  
 CC gene therapy for the treatment of these diseases. The new polypeptide or  
 CC polynucleotide is also useful in diagnostic or research methods. The  
 CC present sequence represents the human tumour endothelial marker 7  
 CC precursor protein

XX SQ Sequence 431 AA;

Query Match 81.2%; Score 2185; DB 6; Length 431;  
 Best Local Similarity 95.2%; Pred. No. 1.1e-200;  
 Matches 416; Conservative 0; Mismatches 7; Indels 14; Gaps 3;

Qy 72 LAMDTLPDNRTRVVEDNHSYVSRLYGSPSPHSELVWDVAEANSQVKIHTILSNTHRQ 131  
 Db 1 LAMDTLPDNRTRVVEDNHSYVSRLYGSPSPHSELVWDVAEANSQVKIHTILSNTHRQ 60  
 Qy 132 ASRVLSDFPFYGHPLRQITATGGPIFMGDVIHRMLTATQYVAPLMANFNPGYSONST 191  
 Db 61 ASRVLSDFPFYGHPLRQITATGGPIFMGDVIHRMLTATQYVAPLMANFNPGYSONST 120  
 Qy 192 VVFDNGTVFVQWHDVYVLOGWEDKGSFTFQALHHDGRIVFAYKEIPMSVPEISSQHP 251  
 Db 121 VVFDNGTVFVQWHDVYVLOGWEDKGSFTFQALHHDGRIVFAYKEIPMSVPEISSQHP 180  
 Qy 252 VKTGLSDAFMILNPSDPVPSRRRSIFEXHRIELDPKVTSMASAVEFTPLTCLQHRSCD 311  
 Db 181 VKTGLSDAFMILNPSDPVPSRRRSIFEXHRIELDPKVTSMASAVEFTPLTCLQHRSCD 240  
 Qy 312 ACMSSDLTFNCSCWCHVLQRCSSGDFDRYRQEW-MDYGCAQAEAGRMCDQD-----ED 363  
 Db 241 ACMSSDLTFNCSCWCHVLQRCSSGDFDRYRQEW-MDYGCAQAEAGRMCDQD-----QDVRLPGMRT 294  
 Qy 364 HDSASPDTSFSPYDGLTFTSSSLFIDSLTDDTKLNPYAGDGLQNNLSPTKGTGPVH 423  
 Db 295 TTSASPDTSFSPYDGLTFTSSSLFIDSLTDDTKLNPYAGDGLQNNLSPTKGTGPVH 354  
 Qy 424 LGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHSTYAE 483  
 Db 355 LGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHSTYAE 414  
 Qy 484 VEPSGHEKEGFMEAEQC 500  
 Db 415 VEPSGHEKEGFMEAEQC 431

Search completed: November 3, 2005, 20:48:47  
 Job time : 169 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 20:46:00 ; Search time 168 Seconds  
(without alignments)

1245.266 Million cell updates/sec

Title: US-09-918-715-230

Perfect score: 2691

Sequence: 1 MRGELMLLVLRRAARALS.....YAEVPSGHEKEGFMEAEQC 500

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867879 seqs, 418409474 residues

Total number of hits satisfying chosen parameters: 1867879

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/PCT PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT NEW PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06 NEW PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06 PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07 NEW PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08 NEW PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08 PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09 NEW PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US10F PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US11A PUBCOMB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US11 NEW PUB.pep.\*  
21: /cgn2\_6/ptodata/1/pubpaa/US60 NEW PUB.pep.\*  
22: /cgn2\_6/ptodata/1/pubpaa/US60 PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2691	100.0	500	10	US-09-918-715-230
2	2691	100.0	500	15	US-10-435-696-79
3	2691	100.0	500	16	US-10-474-794-230
4	2691	100.0	500	16	US-10-357-819-2
5	2691	100.0	500	18	US-10-979-159-230
6	2691	100.0	1002	10	US-09-918-715-179
7	2691	100.0	1002	16	US-10-474-794-179
8	2691	100.0	1002	16	US-10-979-159-179
9	2602	96.7	488	16	US-10-357-819-4
10	2566	95.4	502	14	US-10-156-487A-5
11	2209	82.1	500	10	US-09-918-715-192
					Sequence 230, App
					Sequence 79, Appl
					Sequence 230, App
					Sequence 2, Appl1
					Sequence 179, App
					Sequence 179, App
					Sequence 179, App
					Sequence 4, Appl1
					Sequence 5, Appl1
					Sequence 192, App

12	2209	82.1	500	10	US-09-918-715-297	Sequence 297, App
13	2209	82.1	500	14	US-10-156-487A-6	Sequence 6, Appl1
14	2209	82.1	500	16	US-10-474-794-192	Sequence 192, App
15	2209	82.1	500	16	US-10-474-794-297	Sequence 297, App
16	2209	82.1	500	18	US-10-979-159-192	Sequence 192, App
17	2209	82.1	500	18	US-10-979-159-297	Sequence 297, App
18	2185	81.2	431	10	US-09-912-935-36	Sequence 36, Appl1
19	2185	81.2	431	15	US-10-168-365-36	Sequence 36, Appl1
20	1295	48.1	499	10	US-09-912-935-31	Sequence 31, Appl1
21	1295	48.1	499	15	US-10-168-365-31	Sequence 31, Appl1
22	1295	48.1	529	10	US-09-918-715-189	Sequence 189, App
23	1295	48.1	529	10	US-09-918-715-200	Sequence 200, App
24	1295	48.1	529	10	US-09-912-935-28	Sequence 28, Appl1
25	1295	48.1	529	10	US-09-912-935-40	Sequence 40, Appl1
26	1295	48.1	529	14	US-10-156-487A-4	Sequence 4, Appl1
27	1295	48.1	529	15	US-10-168-365-28	Sequence 28, Appl1
28	1295	48.1	529	16	US-10-474-794-189	Sequence 189, App
29	1295	48.1	529	16	US-10-474-794-200	Sequence 200, App
30	1295	48.1	529	18	US-10-979-159-189	Sequence 189, App
31	1295	48.1	529	18	US-10-979-159-200	Sequence 200, App
32	1294	48.1	529	13	US-10-052-586-472	Sequence 472, App
33	1294	48.1	529	13	US-10-066-500-128	Sequence 128, App
34	1294	48.1	529	14	US-10-174-590-472	Sequence 472, App
35	1294	48.1	529	14	US-10-176-758-472	Sequence 472, App
36	1294	48.1	529	14	US-10-175-737-472	Sequence 472, App
37	1294	48.1	529	14	US-10-174-581-472	Sequence 472, App
38	1294	48.1	529	14	US-10-176-483-472	Sequence 472, App
39	1294	48.1	529	14	US-10-176-749-472	Sequence 472, App
40	1294	48.1	529	14	US-10-176-914-472	Sequence 472, App
41	1294	48.1	529	14	US-10-176-915-472	Sequence 472, App
42	1294	48.1	529	14	US-10-173-706-472	Sequence 472, App
43	1294	48.1	529	14	US-10-175-738-472	Sequence 472, App
44	1294	48.1	529	14	US-10-175-752-472	Sequence 472, App
45	1294	48.1	529	14	US-10-176-482-472	Sequence 472, App

#### ALIGNMENTS

RESULT 1  
US-09-918-715-230  
; Sequence 230, Application US/09918715  
; Publication No. US20030017157A1  
; GENERAL INFORMATION:  
; APPLICANT: Brad St. Croix  
; APPLICANT: Bert Vogelstein  
; APPLICANT: Kenneth Kinzler  
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS  
; FILE REFERENCE: 1107.00134  
; CURRENT APPLICATION NUMBER: US/09/918,715  
; CURRENT FILING DATE: 2001-08-01  
; PRIOR APPLICATION NUMBER: 60/222,599  
; PRIOR FILING DATE: 2000-08-02  
; PRIOR APPLICATION NUMBER: 60/224,360  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/382,850  
; PRIOR FILING DATE: 2000-04-11  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 230  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-918-715-230

Query Match 100.0%; Score 2691; DB 10; Length 500;  
Best Local Similarity 100.0%; Pred. No. 8.3e-249;  
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRGELMLLVLRRAARALSPPQAGHDGPGSGWAAKGTIVRWNRRAARSPGHVSEPR 60  
Db 1 MRGELMLLVLRRAARALSPPQAGHDGPGSGWAAKGTIVRWNRRAARSPGHVSEPR 60

Qy	61	TOLSQDILGGGT	LAMD	TTLPD	NTRV	VEDNHSYV	SRLYG	SPHS	RELW	VDVAE	ANRSQV	120																																								
Db	61	TOLSQDILGGGT	LAMD	TTLPD	NTRV	VEDNHSYV	SRLYG	SPHS	RELW	VDVAE	ANRSQV	120																																								
Qy	121	IHTILSNTH	RQASR	VLVSF	DPPE	YGHPLURQ	TIAT	GGFI	FMG	DVTH	RMILTA	TOYVAPLMA	180																																							
Db	121	IHTILSNTH	RQASR	VLVSF	DPPE	YGHPLURQ	TIAT	GGFI	FMG	DVTH	RMILTA	TOYVAPLMA	180																																							
Qy	181	NFNPGY	SDNSTV	VVYFD	NGTV	FVQW	DHVY	QWDLG	WEDK	SGFT	FQAA	LHHDGR	IVFAYKEIPM	240																																						
Db	181	NFNPGY	SDNSTV	VVYFD	NGTV	FVQW	DHVY	QWDLG	WEDK	SGFT	FQAA	LHHDGR	IVFAYKEIPM	240																																						
Qy	241	SVPEISS	QHPVK	TGLSD	AFML	NPS	PDVPS	RRRS	IF	YHRI	ELD	PSKVT	SMSAVE	TP	300																																					
Db	241	SVPEISS	QHPVK	TGLSD	AFML	NPS	PDVPS	RRRS	IF	YHRI	ELD	PSKVT	SMSAVE	TP	300																																					
Qy	301	LPTCLQ	RHSCD	AC	MS	SDLT	FN	C	SWCH	VLQ	R	CSSG	FDRY	Q	EWMDY	GC	QAE	AG	R	C	B	E	Q	360																												
Db	301	LPTCLQ	RHSCD	AC	MS	SDLT	FN	C	SWCH	VLQ	R	CSSG	FDRY	Q	EWMDY	GC	QAE	AG	R	C	B	E	Q	360																												
Qy	361	DEHD	S	AS	P	D	T	S	F	S	P	Y	G	D	L	T	T	S	S	L	F	I	D	S	L	T	T	E	D	D	T	K	L	N	P	Y	A	G	D	G	I	Q	N	N	L	S	P	K	T	G	T	420
Db	361	DEHD	S	AS	P	D	T	S	F	S	P	Y	G	D	L	T	T	S	S	L	F	I	D	S	L	T	T	E	D	D	T	K	L	N	P	Y	A	G	D	G	I	Q	N	N	L	S	P	K	T	G	T	420
Qy	421	PVHLG	T	T	V	G	I	V	L	A	V	L	L	A	A	I	L	A	G	I	N	G	H	T	P	S	N	A	L	F	F	I	R	R	P	H	H	W	P	A	M	K	P	R	S	H	P	D	K	S	T	480
Db	421	PVHLG	T	T	V	G	I	V	L	A	V	L	L	A	A	I	L	A	G	I	N	G	H	T	P	S	N	A	L	F	F	I	R	R	P	H	H	W	P	A	M	K	P	R	S	H	P	D	K	S	T	480
Qy	481	YAEV	P	S	G	H	E	K	E	G	F	M	E	A	P	Q	500																																			
Db	481	YAEV	P	S	G	H	E	K	E	G	F	M	E	A	P	Q	500																																			

RESULT 2  
US-10-435-696-79  
; Sequence 79, Application US/10435696  
; Publication No. US20040018525A1  
; GENERAL INFORMATION:  
; APPLICANT: Wirtz, Ralph  
; APPLICANT: Munes, Marc  
; APPLICANT: Kallabis, Harald  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS, PROGNOSIS  
; TITLE OF INVENTION: PREVENTION AND TREATMENT OF MALIGNANT NEOPLASIA  
; FILE REFERENCE: Lea 36 108  
; CURRENT APPLICATION NUMBER: US/10/435,696  
; CURRENT FILING DATE: 2003-05-09  
; PRIOR APPLICATION NUMBER: EP03003112.4  
; PRIOR FILING DATE: 2003-02-13  
; PRIOR APPLICATION NUMBER: EP02010291.9  
; PRIOR FILING DATE: 2002-05-21  
; NUMBER OF SEQ ID NOS: 314  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 79  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-435-696-79

[illegible]

```

RESULT 3
US-10-474-794-230
; Sequence 230, Application US/10474794
; Publication No. US20040213793A1
; GENERAL INFORMATION:
; APPLICANT: Carson-Walter, Eleanor
; APPLICANT: St. Croix, Brad
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00179
; CURRENT APPLICATION NUMBER: US/10/474,794
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/308,829
; PRIOR FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 359
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 230
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-474-794-230

```

Db 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPSRRRSIFEYHRIELDSPKVTSMASVEFTP 300  
Qy 301 LPTCLOHRSQDCACMSDDLTFNCSCWCHVLRQCSGFDYRQWMDYGCQAEGRMCEDFQ 360  
Db 301 LPTCLOHRSQDCACMSDDLTFNCSCWCHVLRQCSGFDYRQWMDYGCQAEGRMCEDFQ 360  
Qy 361 DEHDSASPTSPSPYDGLTDTTSSSLFIDSLTTEDDTKLNYPAGDGLQNNLSPKTKGT 420  
Db 361 DEHDSASPTSPSPYDGLTDTTSSSLFIDSLTTEDDTKLNYPAGDGLQNNLSPKTKGT 420  
Qy 421 PVHLGTIVGIVLAVLLVAAILLAGIYINGHTPSNAALFFIERRPHHPAMPKFRSHPDHST 480  
Db 421 PVHLGTIVGIVLAVLLVAAILLAGIYINGHTPSNAALFFIERRPHHPAMPKFRSHPDHST 480  
Qy 481 YAEVPSGHEKEGFMEEAQC 500  
Db 481 YAEVPSGHEKEGFMEEAQC 500

## RESULT 4

US-10-357-819-2  
; Sequence 2, Application US/10357819  
; Publication No. US20040259774A1  
; GENERAL INFORMATION:  
; APPLICANT: Alvarez, Enrique  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Gangolli, Esha A.  
; APPLICANT: Gerlach, Valerie L.  
; APPLICANT: Gorman, Linda  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Li, Li  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Zhong, Mei  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-538A  
; CURRENT APPLICATION NUMBER: US/10/357,819  
; PRIOR FILING DATE: 2003-02-03  
; PRIOR APPLICATION NUMBER: 09/520,781  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 09/584,411  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 09/783,436  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 10/085,198  
; PRIOR FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/353,301  
; PRIOR FILING DATE: 2002-02-01  
; PRIOR APPLICATION NUMBER: 60/355,099  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: 60/356,424  
; PRIOR FILING DATE: 2002-02-12  
; PRIOR APPLICATION NUMBER: 60/358,239  
; PRIOR FILING DATE: 2002-02-20  
; PRIOR APPLICATION NUMBER: 60/358,608  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: 60/359,367  
; PRIOR FILING DATE: 2002-02-25  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 2  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Homo sapiens

## US-10-357-819-2

Query Match 100.0%; Score 2691; DB 16; Length 500;  
Best Local Similarity 100.0%; Pred. No. 8.3e-249;  
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MRGELMLLVLAAREARALSPQAGHDEGPGSGWAAKGTVRGNRRRARESPGHVSEPDFR 60  
Db 1 MRGELMLLVLAAREARALSPQAGHDEGPGSGWAAKGTVRGNRRRARESPGHVSEPDFR 60  
Qy 61 TQSLQDLGGTGLAMDTLPNRTVRVEDNHSYVYRSLYGPSEPHSRRLVWDAENRSQVK 120  
Db 61 TQSLQDLGGTGLAMDTLPNRTVRVEDNHSYVYRSLYGPSEPHSRRLVWDAENRSQVK 120  
Qy 121 IHTILSNTHRQASRVVLSDFPFYGHPLRQITATGCFIEMGDVHRLMLTATQYVAPLMA 180  
Db 121 IHTILSNTHRQASRVVLSDFPFYGHPLRQITATGCFIEMGDVHRLMLTATQYVAPLMA 180  
Qy 181 NFNPYSDNSTVYVFDNGTVFVQWDHVLQGWEDKGSFTFQAALHHDGRIVFAYKEIPM 240  
Db 181 NFNPYSDNSTVYVFDNGTVFVQWDHVLQGWEDKGSFTFQAALHHDGRIVFAYKEIPM 240  
Qy 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPSRRRSIFEYHRIELDSPKVTSMASVEFTP 300  
Db 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPSRRRSIFEYHRIELDSPKVTSMASVEFTP 300  
Qy 301 LPTCLOHRSQDCACMSDDLTFNCSCWCHVLRQCSGFDYRQWMDYGCQAEGRMCEDFQ 360  
Db 301 LPTCLOHRSQDCACMSDDLTFNCSCWCHVLRQCSGFDYRQWMDYGCQAEGRMCEDFQ 360  
Qy 361 DEHDSASPTSPSPYDGLTDTTSSSLFIDSLTTEDDTKLNYPAGDGLQNNLSPKTKGT 420  
Db 361 DEHDSASPTSPSPYDGLTDTTSSSLFIDSLTTEDDTKLNYPAGDGLQNNLSPKTKGT 420  
Qy 421 PVHLGTIVGIVLAVLLVAAILLAGIYINGHTPSNAALFFIERRPHHPAMPKFRSHPDHST 480  
Db 421 PVHLGTIVGIVLAVLLVAAILLAGIYINGHTPSNAALFFIERRPHHPAMPKFRSHPDHST 480  
Qy 481 YAEVPSGHEKEGFMEEAQC 500  
Db 481 YAEVPSGHEKEGFMEEAQC 500

## RESULT 5

US-10-979-159-230  
; Sequence 230, Application US/10979159  
; Publication No. US20050142138A1  
; GENERAL INFORMATION:  
; APPLICANT: Brad St. Croix  
; APPLICANT: Bert Vogelstein  
; APPLICANT: Kenneth Kinzler  
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS  
; FILE REFERENCE: 1107.00134  
; CURRENT APPLICATION NUMBER: US/10/979,159  
; CURRENT FILING DATE: 2004-11-03  
; PRIOR APPLICATION NUMBER: US/09/918,715  
; PRIOR FILING DATE: 2003-08-01  
; PRIOR APPLICATION NUMBER: 60/222,599  
; PRIOR FILING DATE: 2000-08-02  
; PRIOR APPLICATION NUMBER: 60/224,360  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/282,850  
; PRIOR FILING DATE: 2000-04-11  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 230  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-979-159-230

Query Match 100.0%; Score 2691; DB 18; Length 500;  
Best Local Similarity 100.0%; Pred. No. 8.3e-249;

	Matches	500;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
Qy	1	MRGELWLLVLV	REAAARALSPQ	GAGHDEGPGS	WAAKGTVRG	WNRRARESPGHV	SEDPDR	60		
Db	1	MRGELWLLVLV	REAAARALSPQ	GAGHDEGPGS	WAAKGTVRG	WNRRARESPGHV	SEDPDR	60		
Qy	61	TQLSQDILGGGT	LAMDITLPDNR	TRVVEDNHSY	YVGRLYG	PSPHSRELWVD	VAEANRSQVK	120		
Db	61	TQLSQDILGGGT	LAMDITLPDNR	TRVVEDNHSY	YVGRLYG	PSPHSRELWVD	VAEANRSQVK	120		
Qy	121	IHTILSNTHRQAS	RVLVSDFPPF	YGHPLRQIT	ATGGFI	MGWDVTHRM	LATQVYAPLMA	180		
Db	121	IHTILSNTHRQAS	RVLVSDFPPF	YGHPLRQIT	ATGGFI	MGWDVTHRM	LATQVYAPLMA	180		
Qy	181	NFNPGYSDNSTV	VYFDNGTVFV	QWDHVVYLG	QWEDKGSFT	FQAAHLHDGR	IVFAYKEIPM	240		
Db	181	NFNPGYSDNSTV	VYFDNGTVFV	QWDHVVYLG	QWEDKGSFT	FQAAHLHDGR	IVFAYKEIPM	240		
Qy	241	SVPEISSQHPVK	TGLSDAFMLNP	SPDVPSSRRS	IFEXHRI	ELDPSKVT	SMSAVEFTP	300		
Db	241	SVPEISSQHPVK	TGLSDAFMLNP	SPDVPSSRRS	IFEXHRI	ELDPSKVT	SMSAVEFTP	300		
Qy	301	LPTCLQHRSCD	ACMSSDLTFNC	SCWHLQRC	SSGFDRYRQ	QWMDYGCQA	QAEGRMCEDFQ	360		
Db	301	LPTCLQHRSCD	ACMSSDLTFNC	SCWHLQRC	SSGFDRYRQ	QWMDYGCQA	QAEGRMCEDFQ	360		
Qy	361	DEHDHSASPT	SFSPYDGLT	TTTSSSLF	IDSLLT	DEDDTKLNP	YAGGDLQNNLS	SPKTKGT	420	
Db	361	DEHDHSASPT	SFSPYDGLT	TTTSSSLF	IDSLLT	DEDDTKLNP	YAGGDLQNNLS	SPKTKGT	420	
Qy	421	PVHLGTTIVG	IVLAVLLVAA	ITLAGIYING	HPTSTNAAL	FFIERRPHH	WPAMKFRSHPDHST	480		
Db	421	PVHLGTTIVG	IVLAVLLVAA	ITLAGIYING	HPTSTNAAL	FFIERRPHH	WPAMKFRSHPDHST	480		
Qy	481	YAEVEPSGHEK	GFWEAEQC	500						
Db	481	YAEVEPSGHEK	GFWEAEQC	500						

```

RESULT 6
US-09-918-715-179
; Sequence 179, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 179
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-715-179

```

	Query Match	100.0%	Score 2693	DB 10	Length 1002
	Best Local	Similarity	100.0%	Pid. No. 2.5e-248	
	Matches 500	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	1	MRGELLLVLLVLR	AAARALSPQCGAGHDEGPGCSGAAKGTVRGNNRARESPGHVSEDR	60	
Db	503	MRGELLLVLLVLR	AAARALSPQCGAGHDEGPGCSGAAKGTVRGNNRARESPGHVSEDR	562	

Qy	61	TQLSQDLGGGTGLAMDTLPDRNTRVVDNHSYTVSRLYGPPSPHSRRELWVDVAEANRSQVK	120
Db	563	TQLSQDLGGGTGLAMDTLPDRNTRVVDNHSYTVSRLYGPPSPHSRRELWVDVAEANRSQVK	622
Qy	121	IHTILSNTHRQASRVLSFDPFPFYGHPLROIITATGGFIIMGDDVIHRMLTATQVVAPLMA	180
Db	623	IHTILSNTHRQASRVLSFDPFPFYGHPLROIITATGGFIIMGDDVIHRMLTATQVVAPLMA	682
Qy	181	NFNPGYSDNSTVYFDNGTVFVQWHDVYLQGWEDKGSFTFQAALHHDGRIVFAYKEIPM	240
Db	683	NFNPGYSDNSTVYFDNGTVFVQWHDVYLQGWEDKGSFTFQAALHHDGRIVFAYKEIPM	742
Qy	241	SVPEISSOHVPKVTGLSDAFMILNPSDPVPESESRPSIFEVHRIELDSPSKVTSMSAVEFTP	300
Db	743	SVPEISSOHVPKVTGLSDAFMILNPSDPVPESESRPSIFEVHRIELDSPSKVTSMSAVEFTP	802
Qy	301	LPTCLQHRSCDACMSSDLTFNCSSCHVLRQCSSGFDYRQEWMDYGCQAEEAGRMCEDFQ	360
Db	803	LPTCLQHRSCDACMSSDLTFNCSSCHVLRQCSSGFDYRQEWMDYGCQAEEAGRMCEDFQ	862
Qy	361	DEDHDSASPTSPSPYQDGLTTTSSSLFIDSLTTEDDTKLNPYAGGDLQNNLSPTKTKGT	420
Db	863	DEDHDSASPTSPSPYQDGLTTTSSSLFIDSLTTEDDTKLNPYAGGDLQNNLSPTKTKGT	922
Qy	421	PVHLGTTVGIVLAVLLVAALILAGIYINGHPTSNAAALFFTIERRPHHPAMKFRSHPDHST	480
Db	923	PVHLGTTVGIVLAVLLVAALILAGIYINGHPTSNAAALFFTIERRPHHPAMKFRSHPDHST	982
Qy	481	YAEVPSGHEKEGFMEEAQC 500	
Db	983	YAEVPSGHEKEGFMEEAQC 1002	
RESULT 7			
US-10-474-794-179			
; Sequence 179, Application US/10474794			
; Publication No. US20040213793A1			
; GENERAL INFORMATION:			
; APPLICANT: Carson-Walter, Eleanor			
; APPLICANT: St. Croix, Brad			
; APPLICANT: Vogelstein, Bert			
; APPLICANT: Kinzler, Kenneth			
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS			
; FILE REFERENCE: 1107.00179			
; CURRENT APPLICATION NUMBER: US/10/474,794			
; CURRENT FILING DATE: 2003-10-14			
; PRIOR APPLICATION NUMBER: 60/282,850			
; PRIOR FILING DATE: 2001-04-11			
; PRIOR APPLICATION NUMBER: 60/308,829			
; PRIOR FILING DATE: 2001-08-01			
; NUMBER OF SEQ ID NOS: 359			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 179			
; LENGTH: 1002			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-474-794-179			

	Query Match	100.0%	Score 2691;	DB 16;	Length 1002;
	Best Local Similarity	100.0%;	Pred. No. 2.5e-248;		
	Matches 500;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	1	MRGELWLLVLVLR	AAARALSPQPCAGHDEGPGS	AAAKGTVRGNRRARSPGHVSE	60
Db	503	MRGELWLLVLVLR	AAARALSPQPCAGHDEGPGS	AAAKGTVRGNRRARSPGHVSE	562
Qy	61	TQLSQDLGGGT	LTAMDTLTPDNRTRVVDENH	SYVSRLYGCPSEPHSRELWVDV	AEANRSQVK 120
Db	563	TQLSQDLGGGT	LTAMDTLTPDNRTRVVDENH	SYVSRLYGCPSEPHSRELWVDV	AEANRSQVK 622
Qy	121	IHTILSNTHRQAS	RWLSFDPFPFYGHPLRQIT	ATGGGFIEMGDVIHRLMTATQV	VAPLMA 180
Db	623	IHTILSNTHRQAS	RWLSFDPFPFYGHPLRQIT	ATGGGFIEMGDVIHRLMTATQV	VAPLMA 682



```
QY 181 NFNPGYSDNSTVVFYQWHDHVLQGWEDKSGFTFOAALHHDGRIVFAYKEIPM 240
DB 583 NFNPGYSDNSTVVFYQWHDHVLQGWEDKSGFTFOAALHHDGRIVFAYKEIPM 742
QY 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPESRRRSIFEYHRIELDSPSKVTSMAVEFTP 300
DB 743 SVPEISSQHPVKTGLSDAFMILNPSDPVPESRRRSIFEYHRIELDSPSKVTSMAVEFTP 802
QY 301 LPTCLOHRS CDACWSSDLTFNCSCWCHVLRQCSSGFDYRQEWMDYGCQAEGRMCEDFQ 360
DB 803 LPTCLOHRS CDACWSSDLTFNCSCWCHVLRQCSSGFDYRQEWMDYGCQAEGRMCEDFQ 862
QY 361 DEHDASPDTSFSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPTKTGT 420
DB 863 DEHDASPDTSFSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPTKTGT 922
QY 421 PVHLGTTIGVILVAVLLVAAILLAGIYINGHPTNSAALFFTIERRPHHPAMKFRSHPDHST 480
DB 923 PVHLGTTIGVILVAVLLVAAILLAGIYINGHPTNSAALFFTIERRPHHPAMKFRSHPDHST 982
QY 481 YAEVPSGHEKEGFMEAEQC 500
DB 983 YAEVPSGHEKEGFMEAEQC 1002

RESULT 8
US-10-979-159-179
; Sequence 179, Application US/10979159
; Publication No. US20050142138A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/10/979,159
; CURRENT FILING DATE: 2004-11-03
; PRIOR FILING DATE: 2004-09-18, 715
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 179
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-979-159-179

Query Match 100.0%; Score 2691; DB 18; Length 1002;
Best Local Similarity 100.0%; Pred. No. 2.5e-248;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGELMLLVLLREARALSPQAGHDEGPGSGWAAGKGTVRGWNRRARESPGHVSEPOR 60
DB 503 MRGELMLLVLLREARALSPQAGHDEGPGSGWAAGKGTVRGWNRRARESPGHVSEPOR 562
QY 61 TQLSQDLGGGTLMADTLPNRTRVVDNHSYVSRLYGSPHSRRLWVDVAEANSQVK 120
DB 563 TQLSQDLGGGTLMADTLPNRTRVVDNHSYVSRLYGSPHSRRLWVDVAEANSQVK 622
QY 121 IHTILNTHRQASRVLSDFPFYGHPLRQITTIATGGFIPMGDVIHRMLTATQYVAPLMA 180
DB 623 IHTILNTHRQASRVLSDFPFYGHPLRQITTIATGGFIPMGDVIHRMLTATQYVAPLMA 682
QY 181 NFNPGYSDNSTVVFYQWHDHVLQGWEDKSGFTFOAALHHDGRIVFAYKEIPM 240
DB 683 NFNPGYSDNSTVVFYQWHDHVLQGWEDKSGFTFOAALHHDGRIVFAYKEIPM 742
```

```
QY 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPESRRRSIFEYHRIELDSPSKVTSMAVEFTP 300
DB 743 SVPEISSQHPVKTGLSDAFMILNPSDPVPESRRRSIFEYHRIELDSPSKVTSMAVEFTP 802
QY 301 LPTCLOHRS CDACWSSDLTFNCSCWCHVLRQCSSGFDYRQEWMDYGCQAEGRMCEDFQ 360
DB 803 LPTCLOHRS CDACWSSDLTFNCSCWCHVLRQCSSGFDYRQEWMDYGCQAEGRMCEDFQ 862
QY 361 DEHDASPDTSFSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPTKTGT 420
DB 863 DEHDASPDTSFSPYDGLTTTSSSLFIDSLTTEDDTKLNYPAGGDLQNNLSPTKTGT 922
QY 421 PVHLGTTIGVILVAVLLVAAILLAGIYINGHPTNSAALFFTIERRPHHPAMKFRSHPDHST 480
DB 923 PVHLGTTIGVILVAVLLVAAILLAGIYINGHPTNSAALFFTIERRPHHPAMKFRSHPDHST 982
QY 481 YAEVPSGHEKEGFMEAEQC 500
DB 983 YAEVPSGHEKEGFMEAEQC 1002

RESULT 9
US-10-357-819-4
; Sequence 4, Application US/10357819
; Publication No. US20040259774A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez, Enrique
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Bsha A.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Ji, Weizhen
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Miller, Charles E.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Rastelli, Luca
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-538A
; CURRENT APPLICATION NUMBER: US/10/357,819
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 09/584,411
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/783,436
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 10/085,198
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/353,301
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/355,099
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/356,424
; PRIOR FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 60/358,239
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,608
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/359,367
; PRIOR FILING DATE: 2002-02-25
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 4
```

```
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-357-819-4

Query Match      96.7%; Score 2602; DB 16; Length 488;
Best Local Similarity 97.4%; Pred. No. 2.8e-240;
Matches 487; Conservative 1; Mismatches 0; Indels 12; Gaps 1;

QY 1 MRGELWLLVLVLRRAAALSPQAGHDEGPGGAAAGTVRGWNRRARESPPGHVSPDR 60
DQ |||||
DB 1 MRGELWLLVLVLRRAAALSPQAGHDEGPGGAAAGTVRGWNRRARESPPGHVSPDR 60
QY 61 TQSLQDLGGGTAMDTLPDNRTRVVEDNHSYVSRLYGPSEPHSRRELWVDVAEANRSQVK 120
DQ |||||
DB 61 TQSLQDLGGGTAMDTLPDNRTRVVEDNHSYVSRLYGPSEPHSRRELWVDVAEANRSQVK 120
QY 121 IHTILSNTHROASRVLSFDPFPYGHPLRQITATGGFIIMGDVIIHRMLTATQYVAPLMA 180
DQ |||||
DB 121 IHTILSNTHROASRVLSFDPFPYGHPLRQITATGGFIIMGDVIIHRMLTATQYVAPLMA 180
QY 181 NFNPYSDNSTVYVFDNGTVFVQWDHVLQGWEDKGSFTFOAALHHDGRIVFAYKEIPM 240
DQ |||||
DB 181 NFNPYSDNSTVYVFDNGTVFVQWDHVLQGWEDKGSFTFOAALHHDGRIVFAYKEIPM 240
QY 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPESSRRSIFFYHRIELDPSPKTSMSAVEFTP 300
DQ |||||
DB 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPESSRRSIFFYHRIELDPSPKTSMSAVEFTP 300
QY 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLRQCSGFDYRQEWMDYCAQAEAGRMCEDF 359
DQ |||||
DB 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLRQCSGFDYRQEWMDYCAQAEAGRMCEDF 359
QY 360 QD-----BDHDSASPTSFSPYDGLTDTTSSSLFIDSLLTTEDDTKLPVAGDGLQNN 412
DQ |||||
DB 360 QD-----BDHDSASPTSFSPYDGLTDTTSSSLFIDSLLTTEDDTKLPVAGDGLQNN 412
QY 413 LSPKTKGTPVHLGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHHPAMKF 472
DQ |||||
DB 413 LSPKTKGTPVHLGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHHPAMKF 472
QY 473 RSPDHSTYAEVPSGHEGFMFAEQC 500
DQ |||||
DB 475 RSPDHSTYAEVPSGHEGFMFAEQC 502

RESULT 11
US-09-918-715-192
; Sequence 192, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; APPLICANT: Brad St. Croix
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; PRIOR FILING DATE: 2001-08-01
; PRIOR FILING DATE: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR FILING DATE: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR FILING DATE: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 192
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-918-715-192

Query Match      82.1%; Score 2209; DB 10; Length 500;
Best Local Similarity 81.6%; Pred. No. 1.5e-202;
Matches 409; Conservative 46; Mismatches 44; Indels 2; Gaps 2;

QY 1 MRGELWLLVLVLRRAAALSPQAGHDEGPGGAAAGTVRGWNRRARESPPGHVSPDR 59
DQ |||||
DB 1 MRGELWLLVLVLRRAAALSPQAGHDEGPGGAAAGTVRGWNRRARESPPGHVSPDR 59
QY 60 RTQSLQDLGGGTAMDTLPDNRTRVVEDNHSYVSRLYGPSEPHSRRELWVDVAEANRSQVK 119
DQ |||||
DB 61 TQSLQDLGGGTAMDTLPDNRTRVVEDNHSYVSRLYGPSEPHSRRELWVDVAEANRSQVK 120

Query Match      95.4%; Score 2566; DB 14; Length 502;
Best Local Similarity 95.9%; Pred. No. 8.3e-237;
Matches 487; Conservative 0; Mismatches 7; Indels 14; Gaps 3;

QY 1 MRGELWLLVLVLRRAAALSPQAGHDEGPGGAAAGTVRGWNRRARESPPGHVSPDR 60
DQ |||||
DB 1 MRGELWLLVLVLRRAAALSPQAGHDEGPGGAAAGTVRGWNRRARESPPGHVSPDR 60
QY 61 TQSLQDLGGGTAMDTLPDNRTRVVEDNHSYVSRLYGPSEPHSRRELWVDVAEANRSQVK 120
DQ |||||
DB 61 TQSLQDLGGGTAMDTLPDNRTRVVEDNHSYVSRLYGPSEPHSRRELWVDVAEANRSQVK 120
QY 121 IHTILSNTHROASRVLSFDPFPYGHPLRQITATGGFIIMGDVIIHRMLTATQYVAPLMA 180
DQ |||||
DB 121 IHTILSNTHROASRVLSFDPFPYGHPLRQITATGGFIIMGDVIIHRMLTATQYVAPLMA 180
QY 181 NFNPYSDNSTVYVFDNGTVFVQWDHVLQGWEDKGSFTFOAALHHDGRIVFAYKEIPM 240
DQ |||||
DB 181 NFNPYSDNSTVYVFDNGTVFVQWDHVLQGWEDKGSFTFOAALHHDGRIVFAYKEIPM 240
QY 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPESSRRSIFFYHRIELDPSPKTSMSAVEFTP 300
DQ |||||
DB 241 SVPEISSQHPVKTGLSDAFMILNPSDPVPESSRRSIFFYHRIELDPSPKTSMSAVEFTP 300
QY 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLRQCSGFDYRQEWMDYCAQAEAGRMCEDFQ 360
DQ |||||
DB 301 LPTCLQHRSCDACMSSDLTFNCSCWCHVLRQCSGFDYRQEWMDYCAQAEAGRMCEDFQ 360
QY 361 DEHDHDSASPTSFSPYDGLTDTTSSSLFIDSLLTTEDDTKLPVAGDGLQNNLSPKTKGT 420
DQ |||||
DB 361 DEHDHDSASPTSFSPYDGLTDTTSSSLFIDSLLTTEDDTKLPVAGDGLQNNLSPKTKGT 420
QY 421 FVHLGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHHPAMKFRSHPHST 480
DQ |||||
DB 409 FVHLGTIVGIVLAVLLVAAILAGIYINGHPTSNAAALFFIERRPHHPAMKFRSHPHST 468
QY 481 YAEVPSGHEGFMFAEQC 500
DQ |||||
DB 469 YAEVPSGHEGFMFAEQC 488

RESULT 10
US-10-156-487A-5
; Sequence 5, Application US/10156487A
; Publication No. US20030092025A1
; GENERAL INFORMATION:
; APPLICANT: Juan, Todd
; APPLICANT: Bass, Michael B.
; APPLICANT: Oliner, John
; TITLE OF INVENTION: Tumor Endothelial Marker 7a Molecules and Uses Thereof
; FILE REFERENCE: 01-072-A
; CURRENT APPLICATION NUMBER: US/10/156,487A
; PRIOR FILING DATE: 2002-09-10
; PRIOR FILING DATE: 60/293,852
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-156-487A-5

Query Match      95.4%; Score 2566; DB 14; Length 502;
Best Local Similarity 95.9%; Pred. No. 8.3e-237;
Matches 487; Conservative 0; Mismatches 7; Indels 14; Gaps 3;

QY 1 MRGELWLLVLVLRRAAALSPQAGHDEGPGGAAAGTVRGWNRRARESPPGHVSPDR 60
DQ |||||
DB 1 MRGELWLLVLVLRRAAALSPQAGHDEGPGGAAAGTVRGWNRRARESPPGHVSPDR 60
QY 61 TQSLQDLGGGTAMDTLPDNRTRVVEDNHSYVSRLYGPSEPHSRRELWVDVAEANRSQVK 119
DQ |||||
DB 61 TQSLQDLGGGTAMDTLPDNRTRVVEDNHSYVSRLYGPSEPHSRRELWVDVAEANRSQVK 120
```



```
Qy 360 QDEHDSASDPTSPVDGDLTTTSSSLFIDSLTTEDDTKLNPNYAGDGLQNNLSPTKKG 419
Db 361 QDDSHYASDPSFSPNGD-STTSSSLFIDSLTTEDDTKLNPNYAGDGLPDHSSPKSKG 419
Qy 420 TPVHLGTIVGVLAVLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHS 479
Db 420 PPVHLGTIVGVLAVLVAAILAGIYISCHPNNAALFFIERRPHWPAMKFRHHPNHS 479
Qy 480 TYAEVPSGHEKEGFMEAEQC 500
Db 480 TYTEVEPSGHEKEGFVEAEQC 500

RESULT 14
US-10-474-794-192
; Sequence 192, Application US/10474794
; Publication No. US20040213793A1
; GENERAL INFORMATION:
; APPLICANT: Carson-Walter, Eleanor
; APPLICANT: St. Croix, Brad
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00179
; CURRENT APPLICATION NUMBER: US/10/474,794
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/308,829
; PRIOR FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 359
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 192
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-474-794-192

Query Match 82.1%; Score 2209; DB 16; Length 500;
Best Local Similarity 81.6%; Pred. No. 1.5e-202;
Matches 409; Conservative 46; Mismatches 44; Indels 2; Gaps 2;

Qy 1 MRGELWLL-VLVLRARALSPQAGHDEGPGSGWAAGTGVRCWNRARSPGHVSEPD 59
Db 1 MRAQLWLLQLLLRGAARALSPATPAGHNEGQDSAWTAKTRQGSRRSPQAQVLKPG 60
Qy 60 RTQLSQDLGGGTLAMDTLPDNRTRVVDNHSYVSRLYGPSEPHSRRLWVDVAEANSQV 119
Db 61 KTQLSQDLGGGSLAIDLTPDNRTRVVDNHNYYVSRVYGPCEKQSQDLWDLAVANRSHV 120
Qy 120 KIHTILSNTHRQASRVVLSFDPFPGHPLRQITITATGGFIFMGDVIHRMLTATQYVAPLM 179
Db 121 KIHRISSSHRQASRVVLSFDPFPGHPLRQITITATGGFIFMGDMLHRMLTATQYVAPLM 180
Qy 180 ANFNPGYSDNSTVYFDPNGTVFVQWDHVLQGWEDKGSFTFOAALHHDGRIVEAYKEIP 239
Db 181 ANFNPGYSDNSTVAIFDPNGTVFVQWDHVLQDREDRGSFTFOAALHRDGRIVEYKEIP 240
Qy 240 MSVPEISSQHPVKTLGLSDAFMILNPSDPVPESSRRRGI FEYHRIELDPKVTMSAVEFT 299
Db 241 MAVLUDISSAQHPVKAGLSDAFMIILNSSPEVPESQRRTI FEYHRVELDSSKITTSAVEFT 300
Qy 300 PLPTCLOHRSCDACMSSDLTFCNSCHVLCRCSSGFDRIYQEWMDYCAQAEAGRMCEDF 359
Db 301 PLPTCLOHQSCDTCVSSNLTFNCNSCHVLCRCSSGFDRIYQEWLTYCAQAEAGKTCEDF 360
Qy 360 QDEHDSASDPTSPVDGDLTTTSSSLFIDSLTTEDDTKLNPNYAGDGLQNNLSPTKKG 419
Db 361 QDDSHYASDPSFSPNGD-STTSSSLFIDSLTTEDDTKLNPNYAGDGLPDHSSPKSKG 419
Qy 420 TPVHLGTIVGVLAVLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHS 479
Db 420 PPVHLGTIVGVLAVLVAAILAGIYISCHPNNAALFFIERRPHWPAMKFRHHPNHS 479
Qy 480 TYAEVPSGHEKEGFMEAEQC 500
Db 480 TYTEVEPSGHEKEGFVEAEQC 500
```

```
Db 420 PPVHLGTIVGVLAVLVAAILAGIYISCHPNNAALFFIERRPHWPAMKFRHHPNHS 479
Qy 480 TYAEVPSGHEKEGFMEAEQC 500
Db 480 TYTEVEPSGHEKEGFVEAEQC 500

RESULT 15
US-10-474-794-297
; Sequence 297, Application US/10474794
; Publication No. US20040213793A1
; GENERAL INFORMATION:
; APPLICANT: Carson-Walter, Eleanor
; APPLICANT: St. Croix, Brad
; APPLICANT: Vogelstein, Bert
; APPLICANT: Kinzler, Kenneth
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00179
; CURRENT APPLICATION NUMBER: US/10/474,794
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/308,829
; PRIOR FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 359
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 297
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Mouse
US-10-474-794-297

Query Match 82.1%; Score 2209; DB 16; Length 500;
Best Local Similarity 81.6%; Pred. No. 1.5e-202;
Matches 409; Conservative 46; Mismatches 44; Indels 2; Gaps 2;

Qy 1 MRGELWLL-VLVLRARALSPQAGHDEGPGSGWAAGTGVRCWNRARSPGHVSEPD 59
Db 1 MRAQLWLLQLLLRGAARALSPATPAGHNEGQDSAWTAKTRQGSRRSPQAQVLKPG 60
Qy 60 RTQLSQDLGGGTLAMDTLPDNRTRVVDNHSYVSRLYGPSEPHSRRLWVDVAEANSQV 119
Db 61 KTQLSQDLGGGSLAIDLTPDNRTRVVDNHNYYVSRVYGPCEKQSQDLWDLAVANRSHV 120
Qy 120 KIHTILSNTHRQASRVVLSFDPFPGHPLRQITITATGGFIFMGDVIHRMLTATQYVAPLM 179
Db 121 KIHRISSSHRQASRVVLSFDPFPGHPLRQITITATGGFIFMGDMLHRMLTATQYVAPLM 180
Qy 180 ANFNPGYSDNSTVYFDPNGTVFVQWDHVLQGWEDKGSFTFOAALHHDGRIVEAYKEIP 239
Db 181 ANFNPGYSDNSTVAIFDPNGTVFVQWDHVLQDREDRGSFTFOAALHRDGRIVEYKEIP 240
Qy 240 MSVPEISSQHPVKTLGLSDAFMILNPSDPVPESSRRRGI FEYHRIELDPKVTMSAVEFT 299
Db 241 MAVLUDISSAQHPVKAGLSDAFMIILNSSPEVPESQRRTI FEYHRVELDSSKITTSAVEFT 300
Qy 300 PLPTCLOHRSCDACMSSDLTFCNSCHVLCRCSSGFDRIYQEWMDYCAQAEAGRMCEDF 359
Db 301 PLPTCLOHQSCDTCVSSNLTFNCNSCHVLCRCSSGFDRIYQEWLTYCAQAEAGKTCEDF 360
Qy 360 QDEHDSASDPTSPVDGDLTTTSSSLFIDSLTTEDDTKLNPNYAGDGLQNNLSPTKKG 419
Db 361 QDDSHYASDPSFSPNGD-STTSSSLFIDSLTTEDDTKLNPNYAGDGLPDHSSPKSKG 419
Qy 420 TPVHLGTIVGVLAVLVAAILAGIYINGHPTSNAAALFFIERRPHWPAMKFRSHPDHS 479
Db 420 PPVHLGTIVGVLAVLVAAILAGIYISCHPNNAALFFIERRPHWPAMKFRHHPNHS 479
Qy 480 TYAEVPSGHEKEGFMEAEQC 500
Db 480 TYTEVEPSGHEKEGFVEAEQC 500
```

Search completed: November 3, 2005, 20:54:14  
Job time : 170 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2005, 20:45:35 ; Search time 41 Seconds  
(without alignments)  
1173.375 Million cell updates/sec

Title: US-09-918-715-230  
Perfect score: 2691  
Sequence: 1 MRGELMLLVLEAPALSL.....YAEVPSGHEKEGFMEAEQC 500  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	611	22.7	476	T19786	hypothetical prote
2	126	4.7	1161	S31213	nidogen precursor
3	115.5	4.3	1568	T09074	semaphorin recepto
4	103.5	3.8	979	C86446	probable cellulose
5	101.5	3.8	743	T09173	EH domain protein
6	101	3.8	733	E86345	hypothetical prote
7	100	3.7	535	S18606	phosphoenolpyruvat
8	99	3.7	1133	T12529	hypothetical prote
9	98	3.6	345	T16074	hypothetical prote
10	98	3.6	1502	S45429	probable membrane
11	97	3.6	708	T83196	NEEDD-4 ORF - mouse
12	96.5	3.6	397	S33415	corticosteroid-bin
13	96.5	3.6	679	T19703	hypothetical prote
14	95.5	3.5	399	S71480	homeotic protein H
15	95.5	3.5	887	T07842	ubiquitin ligase N
16	95.5	3.5	2120	T30243	alpha tectorin - c
17	94	3.5	774	JC7265	neprilysin (EC 3.4
18	93.5	3.5	979	T12349	protein-tyrosine-p
19	93.5	3.5	996	I48721	PTP 35 protein - m
20	93.5	3.5	1977	S54771	sodium channel alp
21	93	3.5	491	AG3506	phosphoenolpyruvat
22	93	3.5	1085	S55352	IrH1 protein - yea
23	92.5	3.4	441	TJ7653	pectate lyase (EC
24	92.5	3.4	852	A85041	probable receptor
25	91.5	3.4	614	A98241	hypothetical prote
26	91.5	3.4	614	T86888	hypothetical prote
27	91.5	3.4	810	P2WNBB	2a protein - broad
28	91.5	3.4	1042	A57534	mucin 5AC (clone L
29	91	3.4	633	S47144	mating type A prot

30	91	3.4	1175	2	S51005	protein-tyrosine-p
31	90.5	3.4	332	2	JN0067	pregnancy-specific
32	90.5	3.4	424	2	S70399	zona pellucida gly
33	90.5	3.4	4848	2	T30289	pristinamycin I sy
34	90	3.3	561	2	AD2581	phosphoenolpyruvat
35	90	3.3	561	2	B97363	phosphoenolpyruvat
36	90	3.3	798	2	T25104	hypothetical prote
37	90	3.3	1176	2	I58345	protein tyrosine p
38	90	3.3	1179	2	T05673	hypothetical prote
39	90	3.3	1189	2	T51491	hypothetical prote
40	90	3.3	1621	2	T30200	protein-tyrosine k
41	90	3.3	2825	2	T14271	Doc4 protein, stre
42	89.5	3.3	424	2	A34595	pregnancy-specific
43	89.5	3.3	1376	2	G00043	osteonidogen - hum
44	89.5	3.3	1762	2	T03222	probable polyketid
45	89.5	3.3	1840	1	CHRTM1	sodium channel pro

ALIGNMENTS

RESULT 1  
T19786  
hypothetical protein C36E8.3 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T19786  
R;Wilkinson, J.; Barlow, K.  
submitted to the EMBL Data Library, August 1994  
A;Reference number: Z19177  
A;Accession: T19786  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-476 <WIL>  
A;Cross-references: UNIPROT:Q18500; EMBL:Z35597; PIDN:CAA84646.1; GSPDB:GN00021; CESP:C3  
A;Experimental source: clone C36E8  
C;Genetics:  
A;Gene: CESP:C36E8.3  
A;Map position: 3  
A;Introns: 23/1; 101/3; 152/3; 206/3; 235/1; 255/3; 320/1; 445/2  
C;Superfamily: Caenorhabditis elegans hypothetical protein C36E8.3

Query Match	22.7%	Score	611;	DB	2;	Length	476;
Best Local Similarity	34.2%;	Pred. No.	5.6e-42;				
Matches	149;	Conservative	65;	Mismatches	160;	Indels	62;
Gaps	15;						
Qy	46	RRARESPGHVSEPDRTQLS-----QDLGGGTAMDTLPDNRTRVVDNHSYVSRLYGPS	100				
Db	62	RTARAAP-----IPKRSLSAQEDEDIDPATATIP--PDVEVKNDMIDHQYQAETFGVD	115				
Qy	101	EPHSRELWVDVAENRSQVKI-----HTILSNTHRQASRVVLSFDFFPYGHPHQITATG	156				
Db	116	GETLKYWINVBOFMKKPKAVGNTSHPLLSQSVRAVAGARLQKFPFYGHKMGNTLTATG	175				
Qy	157	GFIFMGDVIHRMLTATQYVAPLMAFNPNPGYSNSTRVYFDNGTVFVQVQDHRVYLGWEDK	216				
Db	176	GFYIGDHSNHLAAQYAPLMAFNFT--YLNNSNIVYADGGLFVVEWNVQLKEDKDE	234				
Qy	217	GSTFQAALHGDGRIVFAYKBIKPMSPVPEISSQHPVKVTGLSDAFMLN--PSDPVPSRR	274				
Db	235	HSFTFTILHKGDIVFIYKDVYDISNDSANHPVKLGISDAYMFKHNLQAAP--K	291				
Qy	275	RSIFFEYHRLDPSKVTSMASAVEFTPLCLQHRSDACMSDLT-PNCSMCHVLR---	330				
Db	292	RVYIEYHRIEIAAQKIVSNVTWILKAQPTCISPDCTCTNATLPHFNCMLWCHAKSHHG	351				
Qy	331	--CS--SGFDRVQEWMDYCGAQEAGRMCE-DFQEDHDHSAS-PDTSFSPYDGLTTTS	384				
Db	352	PFCTDEAGLHRRQHWFEQNCYQSKALYCDADDEDEYDEYPKSQLMPNGH---TV	408				
Qy	385	SSLFIDSLTTEDDTKLNPYAGDGLQNLNLSPKTKGTFVHLGTTIVGIVLAVLLVAAILAG	444				
Db	409	LPLDADMKMKTDTTSDSDSEWKGHKKEPK-----GG	442				

QY 445 IYINGHP--TSNAALF 458  
Db 443 VATTAPVGTSSQAATF 458

RESULT 2  
nidogen precursor - sea squirt (Halocynthia roretzi)  
N/Alternate names: entactin  
C/Species: Halocynthia roretzi  
C/Date: 30-Sep-1993 #sequence\_revision 27-Feb-1997 #text\_change 09-Jul-2004  
C/Accession: S31213  
R/Nakae, H.; Sugano, M.; Ishimori, Y.; Endo, T.; Obinata, T.  
Eur. J. Biochem. 213, 11-19, 1993  
A/Title: Ascidian entactin/nidogen. Implication of evolution by shuffling two kinds of c  
A/Reference number: S31213; MUID:93238676; PMID:8477687  
A/Accession: S31213  
A/Molecule type: mRNA  
A/Residues: 1-1161 <NAK>  
A/Cross-references: UNIPROT:Q04901; EMBL:D14038; NID:G217363; PIDN:BAA03127.1; PID:G2173  
C/Superfamily: Ascidian nidogen; EGF homology; LDL receptor WYTD-containing repeat homol  
C/Keywords: basement membrane; collagen binding; disulfide bond; duplication; extracellu  
F/1-20/Domain: signal sequence #status predicted <SIG>  
F/21-1161/Product: nidogen #status predicted <MAT>  
F/274-306/Domain: EGF homology <EG1>  
F/560-574/Region: 3-residue repeats (R-P-V)  
F/603-673/Domain: thyroglobulin type I repeat homology <THY1>  
F/686-748/Domain: thyroglobulin type I repeat homology <THY2>  
F/752-819/Domain: thyroglobulin type I repeat homology <THY3>  
F/900-943/Domain: LDL receptor WYTD-containing repeat homology <YW1>  
F/944-986/Domain: LDL receptor WYTD-containing repeat homology <YW2>  
F/987-1031/Domain: LDL receptor WYTD-containing repeat homology <YW3>  
F/1032-1075/Domain: LDL receptor WYTD-containing repeat homology <YW4>  
F/1076-1114/Domain: LDL receptor WYTD-containing repeat homology <YW5>  
F/1129-1158/Domain: EGF homology <EG2>  
F/107,334,360,484/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.7%; Score 126; DB 1; Length 1161;  
Best Local Similarity 23.5%; Pred. No. 0.094;  
Matches 64; Conservative 34; Mismatches 88; Indels 86; Gaps 12;

QY 137 LSPDFPVGHPLROITATGGFIFM-----GDVTHRLMTATQYVAPLMA----- 181  
Db 47 LSPDIVFYDQYDSVTVHTDGFITLNVGADTDGEVL-----LAPFMSDLDTLSG 97

QY 182 --FNPFGYSNSTVYV-----FDNGTVFVQWDHVYLOGWEDKGSFTFQAA 224  
Db 98 DIFREHKDNATIRANTDVREAFIETAGDFNAGSVFVTVWDVKQASRSDGVTFTFQCI 157

QY 225 LHHGDRIVFA---YKEIPMSVPEIS---SSQHPVKTKGLSDA-----FMILNPSP---DV 269  
Db 158 VATDGAATFAIFLYPDQGLAVGENAVKGVNEVTARAGFNDGGREQLLEILSADELLGGDN 217

QY 270 PESRRRRIFEVHRIELDPKSVTSNAVEFTPLPCLQHRSDCADMSDDLTFNCSWCHVLQ 329  
Db 218 AGSGQGMIFQIGGIMFNNDASKESK-----KH-----HVKX 249

QY 330 RCGSGGFRYRQEW-MDYGCAQAE-GRMCBDF 359  
Db 250 TRQSGFQVSEWNEFDNIDLEACGTGCSDF 281

RESULT 3  
T09074  
semaphorin receptor VESPR - human  
C/Species: Homo sapiens (man)  
C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C/Accession: T09074  
R/Comeau, M.R.; Johnson, R.; DuBoise, R.F.; Petersen, M.; Gearing, P.; van den Bos, T.; H  
Immunity 8, 473-482, 1998  
A/Title: A poxvirus-encoded semaphorin induces cytokine production from monocytes and b  
A/Reference number: Z16555; MUID:98246049; PMID:9586637

A/Accession: T09074  
A/Status: preliminary; translated from GB/EMBL/DBDJ  
A/Molecule type: mRNA  
A/Residues: 1-1568 <COM>  
A/Cross-references: UNIPROT:O60486; EMBL:AF030339; NID:G3176761; PIDN:AAC18823.1; PID:G3  
A/Experimental source: tissue type foreskin; cell type fibroblast  
C/Genetics:  
A/Gene: VESPR  
C/Keywords: receptor; signal transduction

Query Match 4.3%; Score 115.5; DB 2; Length 1568;  
Best Local Similarity 20.6%; Pred. No. 1;  
Matches 77; Conservative 42; Mismatches 118; Indels 137; Gaps 20;

QY 39 GTVRGMNRRARESPGHVSEPDRT-----QLSQDLGGGTFLAMDTTLPDNRTRVVEDNHSYVVS 94  
Db 256 GAATGWPSMAR-----IAQSTEVLFQQAQLDCGHG-----PDGR-----R 292

QY 95 RLYGPSEPHSRELWVDV--AEANRSQVKIHTI-----LSNTHROASRVVLSF----- 139  
Db 293 LLLSSSLVEALDVWAGVFSAAAGEGQERRSPFTTALCLFRMSEIQARAKRVSWDFKTAES 352

QY 140 -----DPPFVGHPLRQITTIATGGFIFMGDVTHRLMTATQYVAPLMAFNFGYSNSTVYV 194  
Db 353 HCEGQOPERVQVFIASSTL-----IHSDLISV-YGTVM-----NRTVLF 391

QY 195 FDNGTVFVQWDHVYLOGWEDKGSFTFQAALHHDGRI--VFAYKEIPMSVPEI---SSSQ 249  
Db 392 LGTG-----DGQLLKVLGENLTSCNCPVIYEIKEE 422

QY 250 HPVKTGLSDAFMLNPSDPVPSERRRSIFVHRIELDPKSVTSNAVEFTPLPCLQHRSS 309  
Db 423 TPV-----FYKLVPPD-----VKNIIYI-----LTAGKEVRRIRVANCNKHS 460

QY 310 CDACMSDLTFNCSWCHVLQRCSSGDFRYRQE---WMDYGCAQAEGRMCEDFQ-----D 361  
Db 461 CSECLTA-TDPHCGWCHSLQRCFTFGDCVHSENLENWLDI-----SSGAKCKPKIQTIRSS 515

QY 362 EDHSASPDTSFSP 375  
Db 516 KEXTVTWVGGSFSP 529

RESULT 4  
C86446  
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: C86446  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: C86446  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-979 <STO>  
A/Cross-references: UNIPROT:Q9FVR3; GB:AE005172; NID:g10801364; PIDN:AAG23436.1; GSPDB:G  
C/Genetics:  
A/Map position: 1

Query Match 3.8%; Score 103.5; DB 2; Length 979;  
Best Local Similarity 21.2%; Pred. No. 5.2;  
Matches 87; Conservative 48; Mismatches 167; Indels 109; Gaps 19;

QY 4 ELWLLVILREARALSPQAGHDSGPGSWAAKGTVRCWNRERESPQHVSEPDRTQL 63



Db 342 EKWKHWKVEDQIKPRPAL---VAPKATWMDGT--HWPGTWAVSGPHSRGDHASV 396  
QY 64 SODL-----GGTLAMD-----TLPDNRFTVEDNHSYVYSLYGSSEPHSR 105  
Db 397 IQVLDPDPGPEVGKGEGRALDLEGVDIRLP-----MLVYSREKRGYDHNK 446  
QY 106 ELWVDVAEANRSQVKIHTILNTHROASRVVLSFDFFPGYHPLQITATGGPIFMGBVI 165  
Db 447 K-----AGAMNALVRASAIMSN-----GPFILNLDCHYVNSRAF-----RDGICFWMDDH 493  
QY 166 HRLMTATQY-----VAPLMAFNPGYSD-----NRTVYV-F 195  
Db 494 GDRVSYVQFPQRFEGIDPDSRYANKNTVFEDINLRALDGIQPMYVGTGCLFRRTALYGF 553  
QY 196 DNGTVFVQVMDHYLVQWEDKSGFTFOALHHDGRIVFAYKEIPMSV---PEISSQHP- 251  
Db 554 NPPDVVFE-----EPPSGSYCFP-----LIRKSPATVASEPPEYITDEDR 595  
QY 252 -----VKTGLSDAFMLNPSDPVPSRRRSIFEYH--RIELDPKVT-SMSAVEFTPLPT 303  
Db 596 FDIGLIRKQFGSSMLYN-SVKVAEFEGRPLATVHSSRLGPPGSLTGRKPLDFATVNE 654  
QY 304 CLQRSCDACMSSDLTNCWSCH--VLQRCSSGFDYRQEMWYGCQAE 352  
Db 655 AVNVISCWYEDKTEWGFNGVNGIYGVTEDEVVTPGRMEKGRSPYCVTEPD 705

## RESULT 5

T09173  
EH domain protein Repsl - mouse  
N/Alternate names: RalBP1-associated EH domain protein Repsl  
C/Species: Mus musculus (house mouse)  
C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C/Accession: T09173  
R/Yamaguchi, A.; Urano, T.; Goi, T.; Feig, L.A.  
J. Biol. Chem. 272, 31230-31234, 1997  
A/Title: An eps homology (EH) domain protein that binds to the ral-GTPase target, RalBP1  
A/Reference number: 216602; MUID:98058900; PMID:9395447  
A/Accession: T09173  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-743 <YAM>  
A/Cross-references: UNIPROT:O54916; EMBL:AF031939; NID:G2677842; PIDN:AA894736.1; PID:G2677842  
A/Experimental source: cell line: C2C12; tissue type: muscle  
C/Genetics:  
A/Gene: repsl  
C/Keywords: signal transduction

Query Match 3.8%; Score 101.5; DB 2; Length 743;  
Best Local Similarity 18.3%; Pred. No. 5.1;  
Matches 91; Conservative 63; Mismatches 166; Indels 177; Gaps 23;  
QY 19 LSPQPGAGH-DEGPGGWAAK-----GTVRGMNRRARESPGHV 55  
Db 78 IPPPGGRGVKKGPGSHDAVQPRPSAQBPSPVSPQSPPTSPHTWKRKSRHPSGN 137  
QY 56 SEPDRQLSQ-----DLGGTLAMDLPDNRFTVEDNHSYVYSLYGSSEPHSR 109  
Db 138 SERPLTGPFPWSPFGDAQGSSAGDAV-----WSGQSPPPQDNWV 179  
QY 110 DVAB-----ANRSQVKIHTILS-NTHROASRVVLSFDFFPGYHPLQITAT 155  
Db 180 SPADTPPTSALLTHWPASVQDQTVTRIVASAATANEIRROSSYEDPW-----KITDE- 232  
QY 156 GGFIFMGDVTHRLMTATQY--VAPLMAFNPGYSDNSTVYVFNQVGVQWHDVY-LQG 212  
Db 233 -----QROYVNVQFKTIQDNLNGFIPG---SAAKEFFTKSKLPILSLSHWELSD 279  
QY 213 WEDKGSFT---FOALHHDGRIVFAYK---EIPMSVPEISSQHPVKTGLSDAFMLNPS 266  
Db 280 FDKDGAULTLDFCAAHF-----LWARKNGYDLPEKLP-----SLMPKLIIDLEDS 325

QY 267 PDVPESRRRSIFEYHRIELDPKSVTSMASAVEFTPLPTCLQHRSCDACMSSDLTNCWSCH 326  
Db 326 ADVGEQGEVGYSGSPAEPSPKSPSPSL-----NQTWPE 361  
QY 327 VLQRCSSGFDYRQEMWYGCQAEGRMCEDFODEHDSASPDTS----- 372  
Db 362 LNQS-----SEWETFS-ERSSSQTLTQF-DSNIAPADPDTAIVHPVPIRMTFSKI 411  
QY 373 -----FSPYDGLTITSSSIFDLSLTEDDTKLPN-----YAGDGLQNNLS 414  
Db 412 HMQEMELKRTSSDHTNTPSLLVKPSDLSENKINSSVKFSGNTVDGYSDDSPPSD-- 469  
QY 415 PKTKGTEV-----HLGT 426  
Db 470 PEQIGSSVTRQSRHSGT 486

## RESULT 6

E86345  
hypothetical protein F16F4.9 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: E86345  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: E86345  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-733 <STO>  
A/Cross-references: UNIPROT:Q91LMN7; GB:AB005172; NID:G8920637; PIDN:AAF81359.1; GSPDB:GN C/Genetics:  
A/Map position: 1  
C/Superfamily: Arabidopsis probable serine/threonine-specific protein kinase PRO25; prot

Query Match 3.8%; Score 101; DB 2; Length 733;  
Best Local Similarity 24.1%; Pred. No. 5.5;  
Matches 59; Conservative 24; Mismatches 72; Indels 90; Gaps 13;

QY 257 SDAFMILNPSDPVPSRRRSI-----FHYHRIELDPKSVTSMASAVE-FTPLPTCL 305  
Db 139 STGMSLCDTPPPNPKNGVGCRTSVSLPLDSHRIETQPSRFENMTSVEHFNPD---- 193  
QY 306 QHRSCD-ACMSSDLTNCFS-----WCHVLQRC-----SS 333  
Db 194 ----CSYAFPEVDGMFNFSSLEDLKLNRVTRFPVLLDWSIGNQTCQVYVGRNICGGNST 249  
QY 334 GFDYRQEMWYGCQAEGRM-----CEDFQD---EDHDSASPD-----SP---SPY 376  
Db 250 CFPDSTGKGVNCKLQGFQDGNPYLSDCCQDINECTTRIHCSDTSTCENTILGSHCCQPS 309  
QY 377 DGLTITSSSFLDLSLTEDDTKLPYAGDGLQNNLSPKTKG-TPVHLGTIVGIVLAVL 435  
Db 310 GSDLNTTWSG-IDTPKEE-----PKYLGWTTVLLGTGTLILL 349  
QY 436 LVAAL 440  
Db 350 TISYI 354

## RESULT 7

S18606  
phenolpyruvate carboxylase (ATP) (EC 4.1.1.49) - Rhizobium sp.  
C/Species: Rhizobium sp.  
C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 27-Oct-2003

C;Accession: S18606  
R;Osteras, M.; Finan, T.M.; Stanley, J.  
Mol. Gen. Genet. 230, 257-269, 1991  
A;Title: Site-directed mutagenesis and DNA sequence of pckA of Rhizobium NGR234, encoding  
A;Reference number: S18606; MUID:92079905; PMID:1720862  
A;Accession: S18606  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-535 <OST>  
A;Cross-references: EMBL:X63291  
C;Superfamily: phosphoenolpyruvate carboxykinase [ATP]  
C;Keywords: carbon-carbon lyase; carboxy-lyase; nucleotide binding; P-loop  
P;234-241/Region: nucleotide-binding motif A (P-loop)

Query Match 3.7%; Score 100; DB 2; Length 535;  
Best Local Similarity 19.8%; Pred. No. 4.3;  
Matches 71; Conservative 50; Mismatches 157; Indels 80; Gaps 15;

Qy 11 VLREAAALSPQAGHDEGPGSCWAAGKGVVGRWNRRAESPGHV--SEPDRTQLSODLG 68  
Db 17 VFRPLGSLQLRAAELEYE-----AFARRGALTGAHGCALCARTGQHTGRSPKDKYVVRDAAT 72  
Qy 69 GGTLAMDT-----LPDNRTRVVED-----NHSYVVSRLYGPSEPHSRELWVDVAEANKRSQ 118  
Db 73 GDQLWWDNNSAISPENFERLRQDMLAHAKGMSLYVQDLVGAGQSGK-----CVADARRHR 127  
Qy 119 VKIHTI-LSNTHROASRVLS-----FDPP-FYGHF-----LRQITIA 154  
Db 128 IRWHSLEIRNLLIRPPREGLASFLPKLTIIDLPSEKANKPERHGCGETIITACDLTKGLVL 187  
Qy 155 TGGFIFMGDVIHRLMTATQYAP-----LMAFNFGYSDNSTVYVF----- 195  
Db 188 IGGTSYAGEMKKSFTVNLNLLPNKAVMPHCSANVGPA---GDTAIFFLSGTGKTTLS 244  
Qy 196 -DNGTVFVQWDHVVYLOQWEDKGSFTFOAALLHDGRIVFAYKEIPMSVPEISSQHPVK 254  
Db 245 ADPNRTILIGDEH---GWSDKGVNPEG-----GCYAKAIRLSEAAPEIPATRRFGT 295  
Qy 255 GLSDAFMLNPPSPVPSRRRSIFEYHRIELDPKSVTMSGAVEFTPTLCLQHRSCDA 312  
Db 296 VMENVLDERRAPDFDNG---SLTENTRIAYPLDFIPNASBTGTAPQPTIIMLTADA 350

RESULT 8  
T12529  
hypothetical protein DKFp3434P113.1 - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 09-Jul-2004  
C;Accession: T12529  
R;Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z17524  
A;Accession: T12529  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1133 <WAM>  
A;Cross-references: UNIPROT:Q9UG37; EMBL:AL080145  
A;Experimental source: adult testis; clone DKFp3434P113  
C;Genetics:  
A;Note: DKFp3434P113.1

Query Match 3.7%; Score 99; DB 2; Length 1133;  
Best Local Similarity 23.0%; Pred. No. 15;  
Matches 67; Conservative 37; Mismatches 119; Indels 68; Gaps 16;

Qy 162 GDVTHRLMTATQYVAPLMAFNPNPGYSNSTVYFDNGTVFVQWDHVVLOQ--WE----- 214  
Db 823 GDIAH--IYDIQTGNKLLTLENPLDANN-----YKNCATFNPTDLDVNDGLVMDVRSQA 876  
Qy 215 -----DKGSFTFOAALLHDGRIVFAYKEI-----PMSVPEISSSQ---HPVKTLGLSD 258  
Db 877 AIHKFDKFNWNVISGVFHPNGLEVIINTEIMDLRTFLLHLHTVPALDQCRVVFNHTGTWYMG 936

Qy 259 AFMLNPPSPVPSRRRSIFEYHRIELDPKSVTMSGAVEFTPTLPTC-LQHRSCDACMSSD 317  
Db 937 AMLQADDEDLMBERMKSPF-----GSSFRFTNATDYKPIATIDVKRNIFDLCTD-- 986  
Qy 318 LTFNCSCHVLRQSCSGFDRYRQEWMDYGC-----AQAEAGRMCEDFODEDH 364  
Db 987 -TKDC-YLAVIENQGS-MDALN--MDTVCRLYEVGRQRLAEDEDEDEDEDEDEDEDD 1040  
Qy 365 DSASPDTFSFYDGDLTSTSSSLFIDSLTATDDTKLNPYAGDGLQNNLSP 415  
Db 1041 DEDDDDTD-----DLDELDTDLLEAELEDDN--NENAGEDG-DNDFSP 1082

RESULT 9  
T16074  
hypothetical protein F1488.6 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: T16074  
R;Geisel, C.  
submitted to the EMBL Data Library, June 1995  
A;Description: The sequence of C. elegans cosmid F1488.  
A;Reference number: Z18456  
A;Accession: T16074  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-345 <GBI>  
A;Cross-references: UNIPROT:Q19446; EMBL:U28737; PID:g860723; PIDN:AAA68276  
A;Experimental source: strain Bristol N2  
C;Genetics:  
A;Gene: CESP:F1488.6  
A;Introns: 17/3; 61/2; 93/3; 144/2; 197/1; 218/3; 255/2; 327/3

Query Match 3.6%; Score 98; DB 2; Length 345;  
Best Local Similarity 18.7%; Pred. No. 3.4;  
Matches 50; Conservative 46; Mismatches 95; Indels 76; Gaps 12;

Qy 240 MSVPETSSQHPVKTLGSLDAFMIINPSPDPVPSRRRSIFEYHRIELDPKSVTMSAVEFT 299  
Db 96 ISVSGIATMKCATLGNASDFFLLMTQSSNQPDNRNQN-----TPIQVTS-NTNQWT 145  
Qy 300 PLPTCLQHRSCDACMSSDLTFNCSCWCHVLRQCSGDFRQEWMDYGCQAEAGRMCEDF 359  
Db 146 LTATI--NRKNG--ADQYGFNCN-----EQCST-----VNDYCYTCGSGNGKTCAS- 191  
Qy 360 QDEHDSDASPTSFSPYDGLTSTSSSLFIDSLTDDTKLNPYAGDGLQNNLSPKTKG 419  
Db 192 -----GDVNPFD-DCSYVDHPISITWSP-----NTQCSASAEN 222  
Qy 420 TPVHLGTIVGIVLAVLLVAAIL-----AGIY-----INGHPTSNAALFF 459  
Db 223 TYFWMISFAIITAILAILLLVLELCCGLFTGQSRSESDGDWIVPPTPKANHELVD 282  
Qy 460 IERRPHHPAMKFRSHPDHSTYAEVPE 486  
Db 283 ADITPHL---QYRRRHQDNNSGESTEP 306

RESULT 10  
S45429  
Probable membrane protein YBL079w - yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein YBL0725  
C;Species: Saccharomyces cerevisiae  
C;Date: 09-Jun-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Jul-2004  
C;Accession: S45429; S45820; S45815; S59226  
R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.  
submitted to the EMBL Data Library, May 1994  
A;Description: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces ce  
A;Reference number: S45387  
A;Accession: S45429  
A;Molecule type: DNA  
A;Residues: 1-1502 <OBE>  
A;Cross-references: UNIPROT:P38181; EMBL:X79489; NID:g496661; PID:g496702



C;Keywords: glycoprotein; steroid binding  
F;1-22/Domain: signal sequence #status predicted <SIG>  
F;23-397/Product: corticosteroid-binding globulin #status experimental #MAT>  
F;89,169,217,232,253,320/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 3.6%; Score 96.5; DB 2; Length 397;  
Best Local Similarity 20.2%; Pred. No. 5.5;  
Matches 53; Conservative 47; Mismatches 104; Indels 59; Gaps 11;  
Qy 61 TQLSQDLGGTGLAMDTLPDRNRTRVEDNHSYVRLYGPSEPHSRELWVDVAE-----A 114  
Db 114 TGLEMMGNVWFLQNLKXDSFLADTKHYVESEALTPSKD-----WTKAGEQINNHNK 168  
Qy 115 NRSQVKHTILSNTHROASRVVLSFDP-----PFYGHPLRQITATGCFIPMGDVIHR 167  
Db 169 NKTQKIEHVVDLSDSATILINIFYLKGWKLFPSPENTRE-----EDFYV 216  
Qy 168 MLTATQVAPLMAFNFGYSDNSTV-----YFDNGTVFVQWDHVYLVQGWEDKGSFT 220  
Db 217 NETSTVKPMVQSGNLSYFRDSPAIPQVMQVYVNGTTFIILPD-----QGQMD-----T 268  
Qy 221 FOAALHHD-----GRIVFAYKEIPMSVPEISSQH-----PVKTLGLSDAFMILNPSPDV 269  
Db 269 VVAALNRDITDRGKMLP-RQNNLYIPKFSMSDTPYDLQVLDVADVGIKDLFTQSDFADT 327  
Qy 270 PESRRRSIFYHR--IELDFSKV 290  
Db 328 TKDTPLTLTLVHLKAMQLQDEGNV 350

RESULT 13  
T19703  
hypothetical protein C34C12.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T19703  
R;Kershaw, J.  
submitted to the EMBL Data Library, December 1994  
A;Reference number: Z19166  
A;Accession: T19703  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-679 <WTL>  
A;Cross-references: UNIPROT:Q09495; EMBL:Z46996; PIDN:CAA87102.1; GSPDB:GNO0021; CESP:C3  
A;Map position: 3  
A;Gene: CESP:C34C12.2  
A;Introns: 5/3; 72/3; 125/3; 173/3; 511/3; 605/3

Query Match 3.6%; Score 96.5; DB 2; Length 679;  
Best Local Similarity 23.3%; Pred. No. 12;  
Matches 70; Conservative 36; Mismatches 130; Indels 65; Gaps 12;  
Qy 16 ARALSPQAGHDGPGCGWAAKTVRGWNRARSPGHVSEPDRTQLSQDLGGTGLAMD 75  
Db 337 AKYLPQDPSTSSACYPKPFYESTSSSRKPIITASP-----PPGRTQISQDLNTEVRY- 392  
Qy 76 TLPDRTRVVEDNHSYVRLYGPSEPHSRELWVDVAENRSQ-----VKIHTILSNTHRQA 132  
Db 393 -----VNSGKPFNFS-----SESNRNKLIPGYIKRPEPRYIKPEGFTSASYKAQ 439  
Qy 133 SRVLSFDPFPYHPLRQITAT-----GGFIPMGDVHRLMTATQVAPLMAFNFGY 186  
Db 440 SEGMSSE-----LKTGSATPENSCKSAHFDPDI-----SSTPYKSHVVVESDEMN 486  
Qy 187 SDNSTVYVF-----DNGTVFVQWDHVYLVQGWEDKGSFTFOAALHHDGRIVFAYKEIPMS 241  
Db 487 SSSSTIGFSEKDKONGAL-----GSKSPMPDITATLHN-----IFDSKEYQSS 531  
Qy 242 VPEISSQHPVKTLGLSDAFMILNPSDPVPSRRRSIFYEYHRIELDPSPKVTSMASVETPL 301  
Db 532 SSTTGSSAPPENSKKSDHF-----DMPDISSTLYRSRVE--PISSSSSGSTSTSAAPRYVK 585

Qy 302 P 302  
Db 586 P 586

RESULT 14  
S71480  
homeotic protein Hox B3 - chicken  
C;Species: Gallus gallus (chicken)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C;Accession: S71480  
R;Scotting, P.J.; Rex, M.  
submitted to the EMBL Data Library, August 1993  
A;Reference number: S71480  
A;Accession: S71480  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-399 <SCO>  
A;Cross-references: UNIPROT:P23682; EMBL:X74506; NID:g398704; PIDN:CAA52613.1; PID:g4433  
C;Genetics:  
A;Gene: Hox-B3  
C;Superfamily: homeotic protein Hox B3; homeobox homology  
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation  
F;160-216/Domain: homeobox homology <HOX>

Query Match 3.5%; Score 95.5; DB 1; Length 399;  
Best Local Similarity 20.5%; Pred. No. 6.7;  
Matches 62; Conservative 40; Mismatches 110; Indels 91; Gaps 10;  
Qy 246 SSSQHPVKTLGLSDAFMILNPSDPVPSRRRSIFYEYHRIELDPSPKVTSM-----AVEFT 299  
Db 94 STSSNIPSGSAKVPVRKPTSVQTP-SLTQKQIFPMMKESQNKQKSSSPSTETCSGKT 152  
Qy 300 PLPTCLQHRSCDACMSDLT-----FNCSWCH-----VLQRCSSGDFRYRQEW-----M 343  
Db 153 PGSSASKRARTAYTSAQLVELEKEFHFNRYLCRPRVENANLLNLSEKQIKIWFQNRM 212  
Qy 344 DYCAQAEAGRMCEDFQDEHDSASPTSFSP-----YGDLTSTSSSLFIDSILT 394  
Db 213 KYKDKQSKGM-----GSSSGGSPGTPGPQPMQSSAGFMNALTMSNVDAPSPS 264  
Qy 395 EDDTKLNPYAGGDLGNL-----SPTKTGTPVHLGT 426  
Db 265 LNKPHQYAHVNTYQNPDKALQOKYTNTAPEYDPVHQNGVAYGTFPSMQGSFVYVG- 323  
Qy 427 IVGIVLAVLLVAAILAGIYINGHTSNAALFFIERRPHPHPAMKFRSHPDHSTYAEVPE 486  
Db 324 -----GNVVDLSLTPSGSLYGLNHLPHHOAA-----NMDYSGPPQWPP 361

Qy 487 SGH 489  
Db 362 SQH 364

RESULT 15  
S70642  
ubiquitin ligase Nedd4 - rat (fragment)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S70642  
R;Staub, O.; Dho, S.; Henry, P.C.; Correa, J.; Ishikawa, T.; McGlade, J.; Rotin, D.  
EMBO J. 15, 2371-2380, 1996  
A;Title: WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na(+)  
A;Reference number: S70642; MUID:96221297; PMID:8665844  
A;Accession: S70642  
A;Status: preliminary; nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-887 <STA>  
A;Cross-references: UNIPROT:Q62940; EMBL:U50842; NID:gl293646; PIDN:AAB48949.1; PID:gl29  
C;Genetics:  
A;Gene: Nedd4  
C;Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiqui

F;54-167/Domain: protein kinase C C2 region homology <KC2>  
F;246-283/Domain: WW repeat homology <WW1>  
F;402-439/Domain: WW repeat homology <WW2>  
F;459-496/Domain: WW repeat homology <WW3>  
F;555-881/Domain: ubiquitin-protein ligase homology <UBI>

```
Query Match      3.5%; Score 95.5; DB 2; Length 887;
Best Local Similarity 19.2%; Pred. No. 20;
Matches 93; Conservative 55; Mismatches 155; Indels 181; Gaps 25;

QY 14 EAARALSP-----QPCAG---HDEGPG---SCWAAKGTVRG-----WNR 47
Db 218 DQAELEPGWVLDPDPAATHLQHPPEPSPLPPGWEERQDVLGRYYVNHESRTQWKP 277
QY 48 ARES-----PGHYSEPDRTQLSDGLGGTGLAMDTLPDNRTR-----VVEDNH 89
Db 278 SPEDDLTDDENGDIQLQAHGAFTRRQISDVVG-----PDNHESPENWEIVREDEN 329
QY 90 SYVYSR-LYGPSEPHSRELWVDVAEANRSQVKIHTILSNTHRQASRVVLSFDFPFYGHPL 148
Db 330 TIYSGQAVQSPPSGHP-DVQVRLAE-----ELDTRLT-----MYGNPA 366
QY 149 RQITIATGGFIFMGDVIRMLTATQYVAPLM-----ANFNPGY-----SDNSTVVVFONGTV 200
Db 367 TSQPTVTSNNHSSRGSSQTCIFEEQPTLPVLLPTSGLPFGWEEKQDDRCGRSYYVDHNSK 426
QY 201 FVVQWDHVLYLQ-----GWEDKGSFTFQAALHHDGRIVF 233
Db 427 -TTTWSKPTWQDDPRSKI PAHLRGKTPVDSNDLGLPLPGWEER-----THTDGRVFF 477
QY 234 AYKEIPMSVPEISSQHPVKTGLSDAFMILNPSPDVPESR-RRSIPFYHRIELDPKSVTS 292
Db 478 INHNKKTQWEDPRMQNVAITG-----PAEPYSRDYKRYEPPFRKLKKQ-----522
QY 293 MSAVEFTPLPTCLQHRSCDACMSDLTFNCSCWCHVLRQCSGFDRTYRQEWMDYGCQAE 352
Db 523 -----TDIPNKFEMK-----LFRANILEDVYRR-IMGVRRADFLK 556
QY 353 GRMCEPDFDE---DHDSASPD-----TSFSPYDGLTITSSSLFIDSLTTEDDT-KLN 401
Db 557 ARLWIEFDGEGKGLDYGGVAREWFFLISKEMFNYYG-----LFEYSATEDNYTLQIN 608
QY 402 PYAG 405
Db 609 PNSG 612
```

Search completed: November 3, 2005, 20:51:20  
Job time : 44 secs

**THIS PAGE BLANK (USPTO)**